EVALUATING GEOGRAPHIC CANCER DISPARITIES IN BALTIMORE CITY, MARYLAND

by
Aracelis Torres, MPH

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March, 2016

Abstract:

Research and policy continue to hold cancer as a top priority in the United States due to its public health burden. As the aging population increases, it is an issue that will likely continue to grow in scope. However, there has been less focus on the complex interaction between social conditions of environments and disparate health outcomes. This dissertation's three aims seek to evaluate the geographic distribution of cancer incidence, cancer mortality, and cancer facilities in Baltimore City, MD as well as the subsequent neighborhood-level correlates that are associated with the observed variability.

The first aim utilized a cohort of Baltimore City female residents (n= 4,966) that were diagnosed with breast, cervical, or colorectal cancer between the ages of 21 to 74 years from 2000 to 2010. The Maryland Cancer Registry provided this data along with the residential address of each cancer case at the time of diagnosis, which allowed for the calculation of incidence rates and cancer stage by neighborhood for each of the three cancer sites. From the cluster detection methods utilized, geographic variation was observed for both outcomes, and it varied by cancer site. The community characteristics explaining this variability also depended on the cancer type.

Using a similar methodology, the second aim utilized the deaths (n= 1,765) that had accrued from the previous aim's cohort of Baltimore City female residents that had developed breast, cervical, or colorectal cancer from 2000 to 2010. The residential address at diagnosis for each cohort member was geocoded to help facilitate the

identification of whether certain neighborhoods had higher cancer mortality than others and whether community-level characteristics played a role in that difference.
Similar as to what was observed for cancer incidence, there were clear differences as to what sections of Baltimore City had a greater burden of mortality. The high burden areas shifted depending on which of the three cancer sites was being evaluated. There was also observed variability in terms of what local-level variables were significantly associated with the geographic aggregation of mortality.

The final aim took the approach of evaluating the location of facilities that provided breast and colorectal cancer services. The addresses for these resources were obtained through publicly available data. This information was then used to construct two measures (service rate and service density) that approximated physical access for each neighborhood. The findings demonstrated that most of the screening and treatment facilities were clustered in the downtown area of Baltimore City.

Overall, the dissertation provided evidence that the evaluation of health outcomes and resources should entail an understanding of the social context in which they occur. The initial findings of the three aims prompt additional research to better hone in on the neighborhood drivers of disease as well as to further investigate neighboring communities with distinct risk profiles. The conclusions reached have the potential to serve as actionable items for policy and resource allocation.

Thesis Readers and Final Oral Exam Committee

Carlos Castillo-Salgado
Thesis Advisor
Professor
Department of Epidemiology
Johns Hopkins School of Public Health

Darcy Phelan-Emrick
Thesis Advisor
Assistant Scientist
Department of Epidemiology
Johns Hopkins School of Public Health

Janice Bowie
Associate Professor
Department of Health, Behavior and Society
Johns Hopkins School of Public Health

Keshia Pollack
Associate Professor
Department of Health Policy and Management
Johns Hopkins School of Public Health

Alternate Committee Members
Corrine Joshu
Assistant Professor
Department of Epidemiology
Johns Hopkins School of Public Health

Terrinieka Powell
Assistant Professor
Department of Population, Family and Reproductive Health
Johns Hopkins School of Public Health

Acknowledgements

There is an African word 'Ubuntu' that roughly translates into "I am what I am because of who we all are." I consider this dissertation a testament to the collective work and effort from not only myself but also my advisors, Carlos Castillo-Salgado and Darcy Phelan-Emrick. I have become the researcher I am today because of your unwavering support and mentorship. Darcy was the first faculty member I contacted when I began my doctoral program search. The welcoming personality she exuded in that first email has held steadfast throughout the nearly five years I have now known her. You have always shown a genuine interest in my development as both a student and a person. I will forever be grateful to your availability and encouragement, which have kept me grounded throughout this journey. I came to know Carlos through a chance meeting early on in the program, and it has since been an encounter that has made all the difference in my doctoral trajectory. Carlos has consistently taken the time to expose me to opportunities that have enriched and pushed the limits of my abilities. There was never a problem you believed was too big for me to solve nor skill too complex for me to learn. Your vote of confidence in my capacity to do meaningful work has been an amazing source of reassurance. I hope that your passion for addressing health inequities is reflected in my pursuit of solving problems that truly matter. To my advisors: I hope I have done you both proud and continue to do so moving forward.

I would also like to thank Eliseo Guallar for his willingness to serve on my thesis committee. You made yourself available in providing feedback and brought a new perspective to my dissertation topic.

I also want to extend my appreciation towards all of the individuals that provided helpful insight by participating in the completion of several doctoral program milestones. Aruna Chandran, Elizabeth Platz, and Kala Visvanathan prepared me during the early stages as members of my departmental exam committee. Janice Bowie, Michael Gibbons, Corinne Joshu, Philip Leaf, and Keshia Pollack provided impactful feedback that helped fine-tune the dissertation proposal when they made themselves available as members of my preliminary school-wide exam committee.

The faculty and staff in the Epidemiology Department have always been a great source of information. In particular, I would like to thank Elizabeth Platz and Kala Visvanathan from the cancer epidemiology track. Your efforts within the field of cancer research have served as inspirational beacons. I should also mention the special place I will hold for my fellow cohort members. I could not have asked for a better group of talented and good-hearted individuals to call my peers. I look forward to seeing you guys accomplish great things as you go forth on your own journeys.

I would also like to recognize Mary Cooke. In the short time that we have known each other, you have been incredibly generous with your time. You quickly believed in the importance and relevance of this work. I am both touched and humbled by the effort you put into trying to address the logistical barriers that always seem to present themselves when accessing data. I am grateful that our paths managed to cross each other.

The pursuit of a doctorate is not only a test of resilience but also a significant financial investment. I do not think I could ever put into words the impact the C. Sylvia and Eddie C. Brown Community Health Scholarship Program has had on my life. Through the generous support of this program, I was able to focus on a dissertation topic I was passionate about and that focused on the local health of Baltimore City residents. The intellectual freedom this funding afforded me is something I will never forget. To Mr. and Mrs. Brown: your philanthropic efforts and dedication to Baltimore are things we should all aspire to emulate. There are few donors that demonstrate the same level of involvement and engagement you have shown to the Brown Scholars. When I think of you both, there is a saying that immediately comes to mind. "We make a living by what we get, but we make a life by what we give." In my opinion, you have made a beautiful life by what you have given to those around you and to the city of Baltimore.

I also cannot mention the Brown Scholars without mentioning Bob Blum. As the faculty moderator for this group, you were always at the ready with a listening ear and a compassionate heart. Beyond funding, the Brown Community Health Scholarship Program has also exposed me to an amazing group of doctoral students. I left every single one of our meetings feeling more inspired and motivated to make a difference. Thanks in particular to Stacy Woods and Karina Christiansen who entered the program with me. I have never known two stronger people. You are both a testament to the fact that there is nothing more resilient than the human spirit. You deserve every bit of happiness and success that comes your way. I hope to continue having the privilege of calling you friends long after the completion of our program.

Additional financial support was provided by the Harvey M. Meyerhoff Fellowship in Cancer Prevention, the Jean Coombs Fund, the Miriam E. Brailey Fund, and the AACR Scholar-in-Training Award.

Also, I would like to recognize my family. My parents have supported all of my endeavors and have always made me feel I had someone in my corner. The foundation of my character is thanks in large part to the home you built and the values you have instilled in me. I am also grateful for my brother and sister. I am fortunate to have had two siblings with which to share life's experiences.

Lastly (but not least), I would like to thank Patrick. I am not quite sure you knew exactly what you were getting yourself into when you began taking this journey alongside me. You witnessed firsthand the all-nighters, the never-ending assignments, and the effort it took to bring it all together in this dissertation. Nevertheless, you patiently endured the difficulties that materialized in the form of time constraints and long-distance commutes. You gave me the space to pursue this dream and have expressed nothing but pride and encouragement. Thank you for your endless support!

Table of Contents

List of tables

List of figures

Abbreviations

MCR: Maryland Cancer Registry

CSA: Community Statistical Area

LISA: Local Indicators of Spatial Association

OLS: Ordinary Least Squares

FDA: U.S. Food and Drug Administration

AAAHC: Accreditation Association for Ambulatory Health Care

AAAASF: American Association for Accreditation of Ambulatory Surgery Facilities

JCAHO: Joint Commission on Accreditation for Healthcare Organizations

CHAPTER 1: INTRODUCTION

Cancer burden and disparities in the United States and Baltimore City, Maryland

Cancer continues to be one of the top health priorities in the United States and the second leading cause of death across all ages. It is projected that over 810,170 new cancer cases and 277,280 cancer deaths will occur among females in 2015.[1] Cervical, breast, and colorectal cancers are estimated to account for almost 40% of these incident cases and 25% of the mortality. In the most recent Annual Report to the Nation on cancer trends, females have experienced an annual decline of 2.8% and 2.5% in colorectal and cervical cancer incidence, respectively, from 2002 to 2011 as breast cancer rates remained stagnant.[2] Within the same time period, annual declines in colorectal cancer deaths (2.9%), breast cancer deaths (1.9%), and cervical cancer deaths (1.3%) were observed. When the report stratified females by racial/ethnic group, there were notable cancer disparities, including a statistically significant increase in breast cancer incidence among black females while white females experienced a decline. While there has been progress in reducing cancer incidence and mortality at the national-level, this improvement is not occurring at the same rate for all subgroups of the population.

Shifting the focus to Maryland, there are about 14,000 new cancer cases per year and 5,100 deaths among female residents.[3] Among all Maryland counties, Baltimore City females yielded the fourth-highest number of incident cases (~1,600) and the most cancer deaths (~700) in 2012.[4] This occurred despite it having just the fourth largest population among Maryland counties. Similar to the estimates at the national-level, disparities were also present as cancer incidence and mortality rates varied by racial/ethnic group (**Table 1.1**).[4] In some cases, Baltimore City even exceeded state and national estimates.

The burden of cancer still looms as a major public health issue while drawing attention at various levels. The national trends have resulted in the United States taking an initiative to include cancer-related objectives in its 10-year agenda, Healthy People 2020, such as increasing cancer screening and decreasing cancer incidence as well as mortality.[5] At the state level, Maryland pushed forth its Comprehensive Cancer Control Plan in July 2011, which highlighted cervical, breast, and colorectal cancer as major targets for site-specific research.[3] A focus on cancer disparities is particularly relevant in Baltimore City given that almost 70% of its residents identify themselves as part of a racial/ethnic minority group.[6] With a growing aging population and increasing diversification of communities, cancer disparities research is an important health priority within Maryland, particularly for Baltimore City.

Shift towards multilevel models in disparities research

As the focus on cancer disparities grows, it is closely followed by an increased interest in studying the disease from a broader scope. Researchers have been expanding beyond the traditional medical and etiological models that previously focused on individual-level risk factors. This shift has become readily apparent at the national-level as the Healthy People 2020 agenda specifically mentioned the need to "create social and physical environments that promote good health for all" as one of its four overarching objectives.[5] The Department of Health for the United States has decided to accomplish this goal by focusing on a place-based framework centered around five key areas of social determinants: 1) economic stability; 2) education; 3) social and community context; 4) health and health care; and 5) neighborhood and built environment. This development demonstrated the recognition that residential context affects individual-level risk, and the conditions in which citizens reside partly explain why some individuals are healthier than others.

Baltimore City has a long history of having social determinants unequally distributed across neighborhoods. This inequity became part of a national discussion as riots took over the city in April 2015. These protests brought to light the backdrop of economic and social conditions affecting the city's dynamics, such as concentrated neighborhood poverty and racial segregation.[7] Upon observing an overview of social determinants across Baltimore City, it becomes apparent that multiple indicators quickly accumulate in the same neighborhoods. Communities that have a higher concentration of African-American residents are the same areas that yield the most single female-headed households and families earning less than $25,000 a year (**Figure 1.1**). A place-based framework acknowledges that neighborhood conditions affect health outcomes. Disease does not occur in isolation nor independently from social context.

This broader perspective on disease is being put forth not just by federal organizations. State- and local-level parties, including those within Maryland, have echoed similar stances. The Baltimore Neighborhood Indicators Alliance-Jacob France Institute at the University of Baltimore has been compiling community data to track the quality of life for all Baltimore City Community Statistical Areas (CSAs).[8] The Baltimore City Health Department collected data on major health outcomes for all of these communities and disseminated the information for each of these 55 areas as part of the Neighborhood

Health Profile Report.[9] One of the most striking findings from this 2011 report was the wide gap in life expectancy. As a whole, the average life expectancy in Baltimore City was 71.8 years. However, two neighborhoods, Roland Park and Hollins Market, separated by only six miles had a 20-year gap in life expectancy of 83.1 years versus 64 years, respectively. These data fueled the Maryland Health Improvement and Disparities Reduction Act of 2012, which aims to reduce the geographic variation in health outcomes by establishing Health Enterprise Zones in communities with limited resources.[10]

Geographic disparities resonate within Baltimore City not only in terms of life expectancy but also cancer outcomes. The state's Comprehensive Cancer Control Plan noted the importance of addressing geographic variation in light of Baltimore City possessing a cancer mortality rate 23% higher than the state average in the 2010 Baltimore City Health Disparities Report Card.[3] As health policy attempts to address these differences across neighborhoods, cancer research has also moved towards integrating multilevel models for this purpose as well. Researchers from the Centers for Population Health and Health Disparities reflected on the sources of health disparities and acknowledged early findings demonstrating that neighborhood context has a substantial influence on individual health independent of individual-level risk factors.[11] There has been momentum to "visualize the multiple influences on cancer and cancer disparities and understand the complex ways in which they interact with one another to produce worse outcomes for some groups than others."[12] As a result, there has been growing literature on the utilization of geostatistical methods to study disease variation across regions in more depth. These methods have been implemented throughout the disease spectrum with some studies aimed at identifying the current placement of prevention services and facilities to describe physical access.[13–18] Others have given a closer look at more secondary and tertiary outcomes, such as cancer incidence and survival.[19–24]

Most recently, a large study on colorectal cancer survival used information from over half a million members of the American Association of Retired Persons (AARP) spanning across six states.[25] The findings demonstrated that individuals living in neighborhoods with high levels of socioeconomic deprivation had a higher risk of overall death and cancer-specific death than those living in areas with low levels of deprivation. While these results are in line with a multilevel perspective of disease, the patterns and associations were observed at the census tract-level across various states. These geographic boundaries may not always align with local perception as to what constitutes a neighborhood. Neighborhood-level findings should be disseminated within a framework that reflects a meaningful understanding of the local culture and community affiliations.

Conceptual Framework

In addressing the relationship between population risk and individual risk, the Warnecke Model (**Figure 1.2**) is a multilevel perspective highlighting social context, social relationships, and physical context of neighborhoods as the mechanisms through which institutional factors affect individual-specific cancer risk.[11] This framework presents an understanding that cancer disparities are due to a complex relationship across several levels of factors. Its recentness and the process by which it was developed drove the selection of this particular model. While other paradigms exist that visualize the multi-faceted nature of health, such as the widely used Social Model of Health by Dahlgren and Whitehead or the Brunner and Marmot model of social determinants, the Warnecke Model benefited from having been created more recently.[26,27] The breadth and depth of the landscape of social determinants has changed very quickly within a short span of time. The topic of health disparities and the efforts directed towards addressing them have exponentially gained traction over the last two to three decades. The comprehensiveness of the Warnecke Model reflects a more finely tuned understanding of what and how multilevel features interact with one another to produce disparate health outcomes.

The other reason for selecting the Warnecke Model as the framework behind the dissertation's analyses stemmed from the way in which it was created. The National Institutes of Health sponsored the creation of eight Centers for Population Health and Health Disparities (CPHHDs) with the goal of developing new models and conducting multilevel research that explored social determinants.[11] The Warnecke Model was created through a collective effort and utilized by this network of centers in the development of its respective research questions as well as experiments. Its framework hypothesized the existence of three main types of determinants: 1) distal determinants (policies, social conditions, and institutional context); 2) intermediate determinants (social and physical context of neighborhoods); and 3) proximal determinants (individual-level characteristics). The model demonstrates that there is interaction across these three types of determinants that result in differential health outcomes across different populations.

The subsequent research conducted by the CPHHDs validated the relationships and interaction displayed within the model's framework, which confirmed its credibility.

This dissertation focuses on the analysis of the social and physical context of neighborhoods to explain geographic differences in cancer incidence, cancer mortality, and locations of cancer-related services in Baltimore City, Maryland.

Dissertation goals and specific aims

Overall goal of the dissertation

This dissertation seeks to expand on the current research in two particular ways. First, it aims to assess the influence of residence across the disease spectrum rather than focusing on one particular aspect of the continuum by looking at cancer incidence and mortality as well as the geographic distribution of screening/treatment facilities. Secondly, it focuses on local patterns to assess the association between neighborhood characteristics and disease within the scope of Baltimore City. These associations may not have been apparent in a larger study that spanned across an expansive region, such as the analysis of AARP members from multiple states. As will be described in further detail later, unique neighborhood-level data from the Baltimore Neighborhood Indicators Alliance, which are not routinely collected by other regions, were integrated into the analysis. Once the association between community characteristics and cancer risk is appropriately studied, then the findings can be generalized with more confidence to other communities with similar neighborhood structures as Baltimore City.

Specific Aim 1

Identify neighborhood characteristics associated with the geographic variation of cancer incidence for breast, cervical, and colorectal cancers among females residing in Baltimore City, Maryland who were diagnosed between the ages of 21-74 years from 2000 to 2010.

Hypothesis: Incidence and cancer stage among female cancer cases will exhibit geographic variation that will be correlated and explained by neighborhood characteristics.

Specific Aim 2

Identify neighborhood characteristics associated with the geographic variation of cancer mortality from date of cancer diagnosis to date of cancer-specific mortality for breast, cervical, and colorectal cancers among females residing in Baltimore City, Maryland who were diagnosed between the ages of 21-74 years from 2000 to 2010.

Hypothesis: Cancer mortality of female cancer cases will exhibit geographic variation that will be correlated and explained by neighborhood characteristics.

Specific Aim 3

Identify neighborhood characteristics associated with the geographic variation in breast and colorectal cancer-related services in Baltimore City, Maryland.

Hypothesis: Cancer screening/treatment facilities will exhibit geographic variation that will be correlated and explained by neighborhood characteristics.

Organization of this dissertation

This chapter, Chapter 1, serves as an introduction and overview. Chapters 2, 3, and 4 each consist of a geospatial analysis evaluating the association between the characteristics of Baltimore City neighborhoods and three separate outcomes: cancer incidence (Aim 1); cancer survival (Aim 2); and physical locations of cancer-related centers/services (Aim 3), respectively. Each chapter contains separate appendices. Chapter 5 contains a summary of the results and final conclusions.

Data Sources

<u>Maryland Cancer Registry</u>: The Maryland Cancer Registry (MCR), Center for Cancer Prevention and Control, was used to obtain information on incidence (Aim 1) and vital status (Aim 2) for breast, cervical, and colorectal cancer cases occurring from 2000 to 2010 among the study population of females 21-74 years old residing in Baltimore City. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the National Program of Cancer Registries of the Centers for Disease Control and Prevention for the funds that support the collection and availability of the cancer registry data. The registry has been in place since 1992 and is maintained by the Maryland Department of Health and Mental Hygiene, which also handles the quality assurance components. As required by law, hospitals, freestanding radiation therapy centers, laboratories, and physicians must report tumors to the registry within 6 months of the diagnosis.[28] The Maryland Cancer Registry also has agreements with nearby states (Delaware, Pennsylvania, Virginia, West Virginia, and District of Columbia) to receive information on Maryland residents that are diagnosed or receive cancer treatment outside of

Maryland. Of the information collected as part of the registry's protocol, a *geocoded residential address at the time of cancer diagnosis* was obtained for each cancer case as well as the following individual-level covariates: *age, race/ethnicity, cancer site, cancer grade, tumor size, date of cancer diagnosis, type of treatment, date of death and cause*.

Vital Signs 10 report: This report created by the Baltimore Neighborhood Indicators Alliance-Jacob France Institute at the University of Baltimore was used to obtain information on neighborhood-level characteristics, including U.S. Census data. The organization has collected community data since 2000 to follow the changes in the quality of life for all Baltimore City neighborhoods.[8] To maintain the consistency of geographic boundaries for the tracking of communities, neighborhoods were designated according to Community Statistical Area (CSA) boundaries, which are comprised of a collection of census tracts. All geospatial analyses (Aims 1, 2, and 3) were conducted relative to these neighborhood designations (**Figure 1.3**). The names of these geographic units can be found in **Table 1.2**. Data describing each community were measured according to 7 major areas that were considered central to adequately capturing the quality of life. These areas were: 1) *U.S. Census data*; 2) *housing and community development*; 3) *crime and safety; 4) education and youth; 5) children and family health*; 6) *workforce/economic development*; and 7) *sustainability*. Across these domains, a total of 66 indicators were evaluated.

Baltimore City directories: Since there is currently no measure to describe physical access to cancer screening and treatment within a neighborhood, directories as well as publicly available databases were used to identify facilities that provided these cancer-related services for any of the two cancer sites of interest: 1) breast; and 2) colorectal. All of the agencies were contacted to confirm that *cancer services were provided* as well as the *address at which services were rendered*. These prevention and treatment sites were categorized by neighborhood to construct measures that will be described in Chapter 4 within the methodology section.

References:

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015;65(1):5–29.

2. Kohler BA, Sherman RL, Howlader N, et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. J Natl Cancer Inst 2015;107(6):djv048.

3. Maryland Department of Health & Mental Hygiene. Maryland Comprehensive Cancer Control Plan. Baltimore, MD: Centers for Disease Control and Prevention; 2011.

4. National Institutes of Health. State Cancer Profiles: Quick Profiles (Maryland) [Internet]. National Cancer Institute. Available from: http://statecancerprofiles.cancer.gov/quick-profiles

5. Office of Disease Prevention and Health Promotion. Healthy People 2020: Cancer Objectives [Internet]. U.S. Department of Health and Human Services. Available from: http://www.healthypeople.gov/2020/topics-objectives/topic/cancer

6. Maryland Department of Health & Mental Hygiene Minority Health and Health Disparities. Maryland Chartbook of Minority Health And Minority Health Disparities Data With Sections on Gender-specific Health And Jurisdiction-specific Health. Baltimore, MD: 2009.

7. Good fortune, dire poverty, and inequality in Baltimore: An American story [Internet]. The Brookings Institution. [cited 2016 Jan 31];Available from: http://www.brookings.edu/blogs/the-avenue/posts/2015/05/11-poverty-inequality-baltimore-berube-mcdearman

8. Baltimore Neighborhood Indicators Alliance–Jacob France Institute. Vital Signs 10: Measuring Baltimore's Continued Progress Towards Strong Neighborhoods and a Thriving City. Baltimore, MD: University of Baltimore; 2012.

9. Baltimore City Health Department. Neighborhood Health Profiles [Internet]. 2012;Available from: http://www.baltimorehealth.org/neighborhood.html

10. Maryland General Assembly. Senate Bill 234 [Internet]. 2012;Available from: http://mlis.state.md.us/2012rs/billfile/sb0234.htm.

11. Warnecke RB, Oh A, Breen N, et al. Approaching health disparities from a population perspective: the National Institutes of Health Centers for Population Health and Health Disparities. Am J Public Health 2008;98(9):1608–15.

12. Gehlert S, Colditz GA. Cancer disparities: unmet challenges in the elimination of disparities. Cancer Epidemiol Biomarkers Prev 2011;20(9):1809–14.

13. Lian M, Jeffe DB, Schootman M. Racial and geographic differences in mammography screening in St. Louis City: a multilevel study. J Urban Health 2008;85(5):677–92.

14. Sheehan TJ, Gershman ST, MacDougall LA, et al. Geographic assessment of breast cancer screening by towns, zip codes, and census tracts. J Public Health Manag Pract 2000;6(6):48–57.

15. King SJ, Livingston PM, Turner L, et al. Mapping the information resources available to patients with colorectal cancer. Eur J Cancer Care (Engl) 2010;19(4):492–500.

16. Maheswaran R, Pearson T, Jordan H, Black D. Socioeconomic deprivation, travel distance, location of service, and uptake of breast cancer screening in North Derbyshire, UK. J Epidemiol Community Health 2006;60(3):208–12.

17. Mobley LR, Kuo T-M, Urato M, Subramanian S. Community contextual predictors of endoscopic colorectal cancer screening in the USA: spatial multilevel regression analysis. Int J Health Geogr 2010;9:44.

18. Rahman S, Price JH, Dignan M, Rahman S, Lindquist PS, Jordan TR. Access to Mammography Facilities and Detection of Breast Cancer by Screening Mammography: A GIS Approach. Int J Canc Prev 2009;2(6):403–13.

19. Hsu CE, Mas FS, Hickey JM, Miller JA, Lai D. Surveillance of the colorectal cancer disparities among demographic subgroups: a spatial analysis. South Med J 2006;99(9):949–56.

20. McElroy JA, Remington PL, Gangnon RE, Hariharan L, Andersen LD. Identifying geographic disparities in the early detection of breast cancer using a geographic information system. Prev Chronic Dis 2006;3(1):A10.

21. Mullee MA, De Stavola B, Romanengo M, Coleman MP. Geographical variation in breast cancer survival rates for women diagnosed in England between 1992 and 1994. Br J Cancer 2004;90(11):2153–6.

22. Fairley L, Forman D, West R, Manda S. Spatial variation in prostate cancer survival in the Northern and Yorkshire region of England using Bayesian relative survival smoothing. Br J Cancer 2008;99(11):1786–93.

23. Rushton G, Peleg I, Banerjee A, Smith G, West M. Analyzing geographic patterns of disease incidence: rates of late-stage colorectal cancer in Iowa. J Med Syst 2004;28(3):223–36.

24. Schootman M, Jeffe DB, Lian M, Gillanders WE, Aft R. The role of poverty rate and racial distribution in the geographic clustering of breast cancer survival among older women: a geographic and multilevel analysis. Am J Epidemiol 2009;169(5):554–61.

25. Lian M, Schootman M, Doubeni CA, et al. Geographic variation in colorectal cancer survival and the role of small-area socioeconomic deprivation: a multilevel survival analysis of the NIH-AARP Diet and Health Study Cohort. Am J Epidemiol 2011;174(7):828–38.

26. Dahlgren, G W M. Policies and Strategies to Promote Equity in Health. Stockholm: Institute of Future Studies; 1991.

27. Brunner, E M M. Social Organization, Stress, and Health. Oxford: Oxford University Press; 1999.

28. Department of Mental Health and Hygiene. Maryland Cancer Registry [Internet]. Available from: http://phpa.dhmh.maryland.gov/cancer/SitePages/mcr_home.aspx

Table 1.1 County-, State-, Country-Specific Annual Cancer Incidence and Mortality Rates* for Females by Race, 2008-2012

	Baltimore City		Maryland		United States	
	Whites	**Blacks**	**Whites**	**Blacks**	**Whites**	**Blacks**
All Cancer Sites						
Incidence	487.5	393.2	426.3	379.8	417.9	390.8
Mortality	183.5	196.2	145.5	164.7	145.6	166.3
Breast Cancer						
Incidence	138.3	123.5	130.3	129.1	124.1	121.4
Mortality	23.8	31.5	21.7	30.4	21.3	30.2
Cervical Cancer						
Incidence	10.1	10.3	6.0	8.2	7.5	9.8
Mortality	4.2	5.3	1.7	3.3	2.1	4.0
Colorectal Cancer						
Incidence	40.8	47.4	32.8	39.3	35.7	43.4
Mortality	15.4	21.7	11.4	17.3	12.7	17.8

*Rates are per 100,000

Figure 1.1: Geographic distribution of three social indicators by quintiles across Baltimore City, Maryland

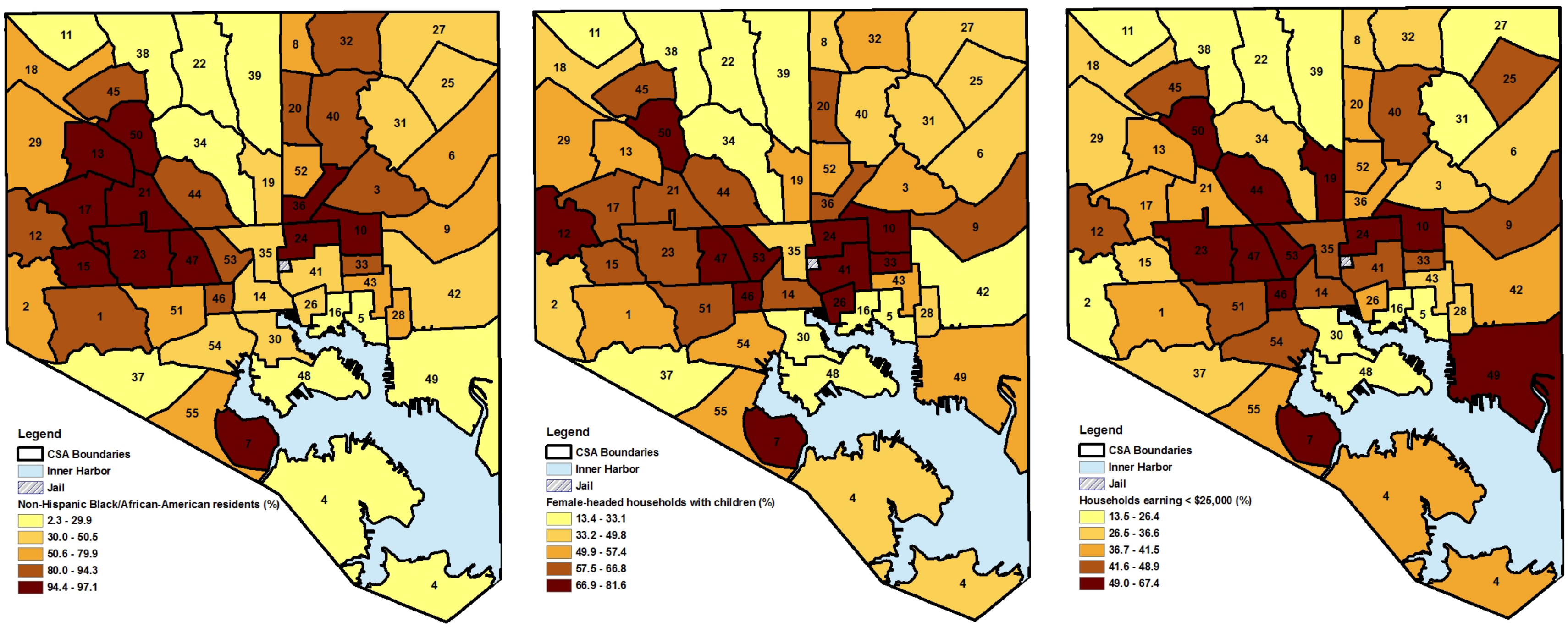


Figure 1.2: Adapted Warnecke model for cancer disparities

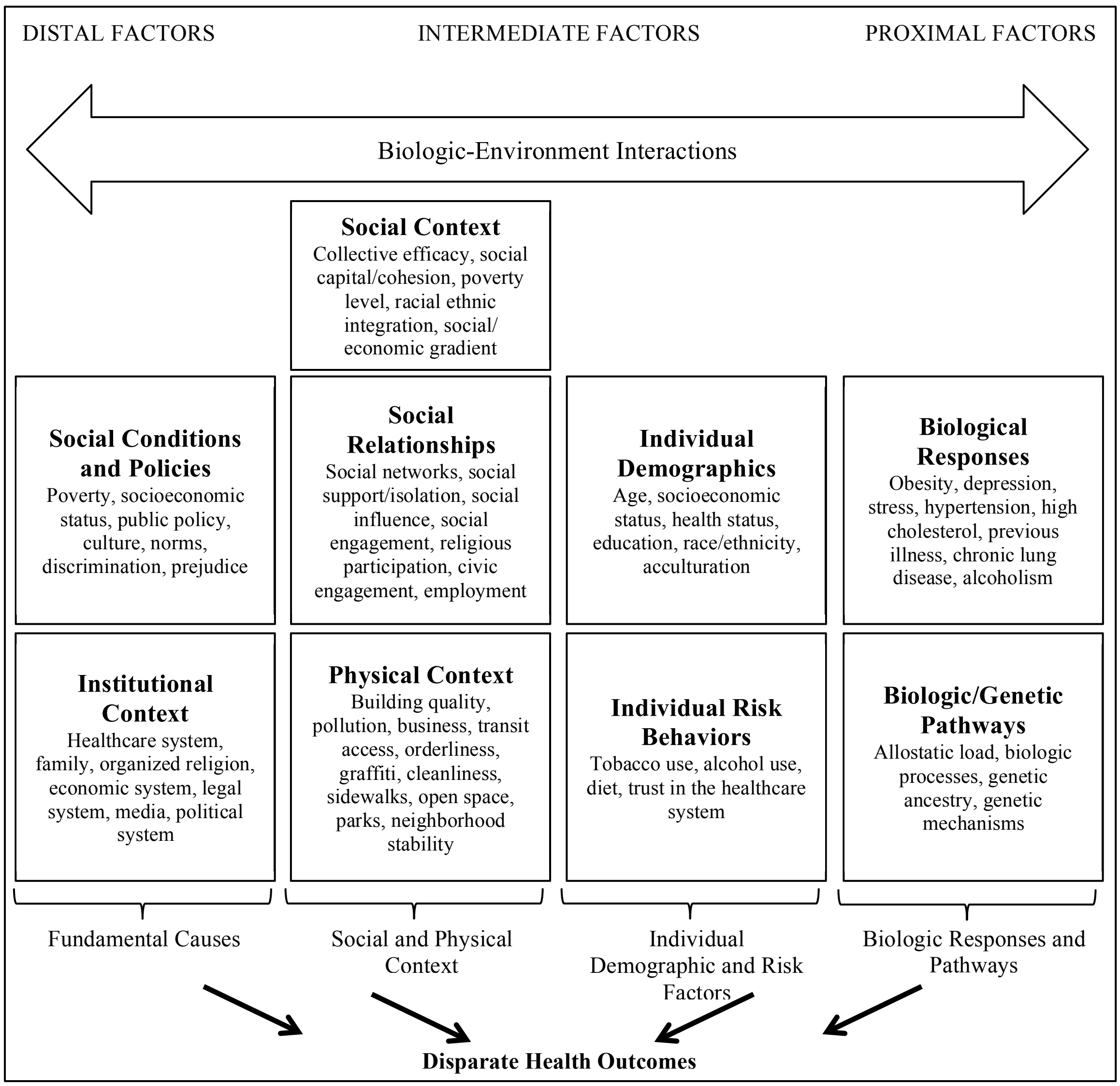


Figure 1.3: Baltimore City Community Statistical Area (CSA) map

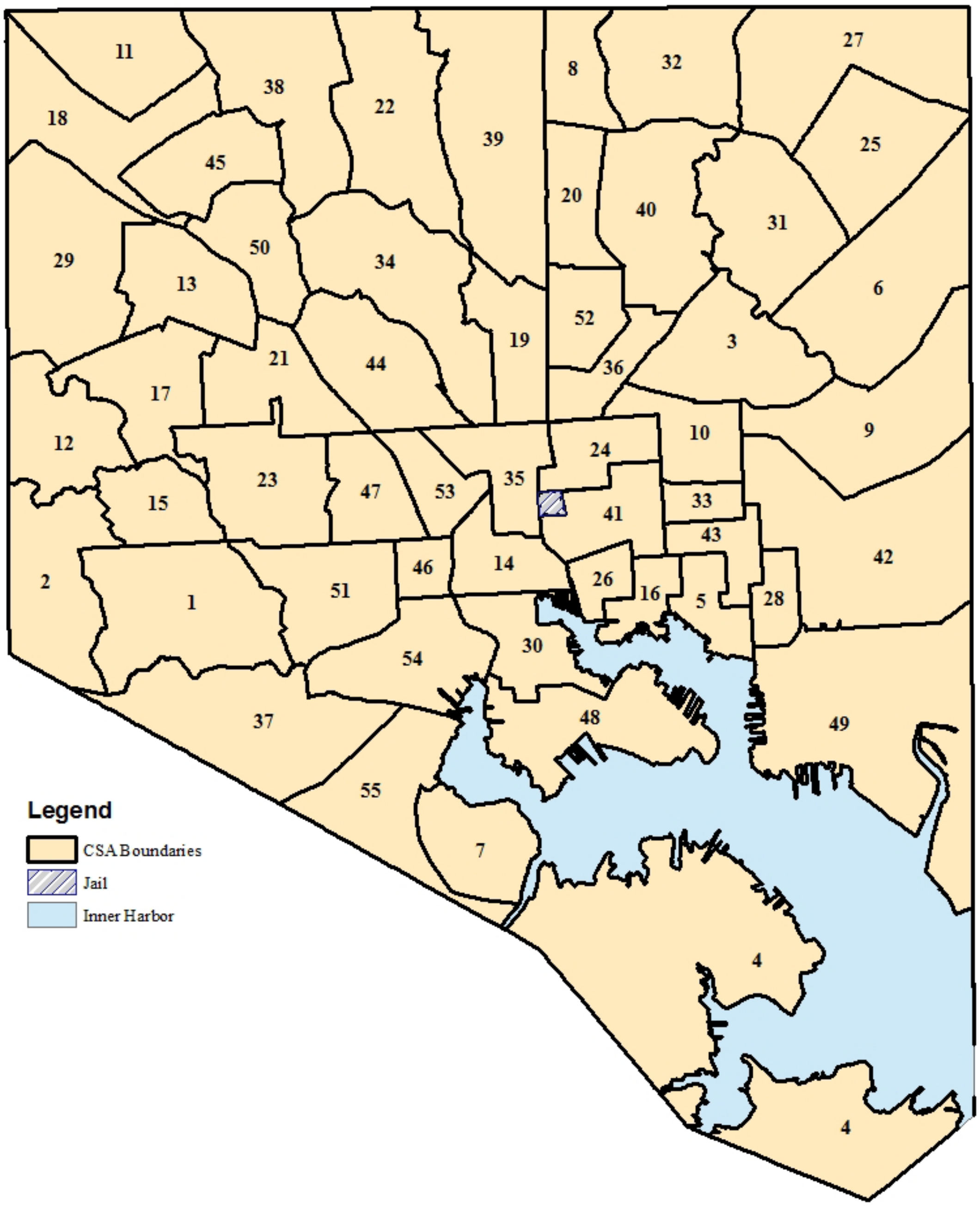


Table 1.2: Key for Community Statistical Area (CSA) map

Allendale/Irvington/S. Hilton	1	Howard Park/West Arlington	29
Beechfield/Ten Hills/West	2	Inner Harbor/Federal Hill	30
Belair-Edison	3	Lauraville	31
Brooklyn/Curtis Bay	4	Loch Raven	32
Canton	5	Madison/East End	33
Cedonia/Frankford	6	Medfield/Hampden/Woodberry/Remington	34
Cherry Hill	7	Midtown	35
Chinquapin Park/Belvedere	8	Midway/Coldstream	36
Claremont/Armistead	9	Morrell Parkk/Violetville	37
Clifton-Berea	10	Mount Washington/Coldspring	38
Cross-Country/Cheswolde	11	North Baltimore/Guilford/Homeland	39
Dickeyville/Franklintown	12	Northwood	40
Dorchester/Ashburton	13	Oldtown/Middle East	41
Downtown/Seton Hill	14	Orangeville/E. Highlandtown	42
Edmonson Village	15	Patterson Park North & East	43
Fells Point	16	Penn North/Reservoir Hill	44
Forest Park/Walbrook	17	Pimlico/Arlington/Hilltop	45
Glen-Fallstaff	18	Poppleton/The Terraces/Hollins Market	46
Greater Charles Village/Barclay	19	Sandtown-Winchester/Harlem Park	47
Greater Govans	20	South Baltimore	48
Greater Mondawmin	21	Southeastern	49
Greater Roland Park/Poplar Hill	22	Southern Park Heights	50
Greater Rosemont	23	Southwest Baltimore	51
Greenmount East	24	The Waverlies	52
Hamilton	25	Upton/Druid Heights	53
Harbor East/Little Italy	26	Washington Village/Pigtown	54
Harford/Echodale	27	Westport/Mount Winans/Lakeland	55
Highlandtown	28		

CHAPTER 2: EVALUATING NEIGHBORHOOD CORRELATES AND GEOGRAPHIC DISTRIBUTION OF BREAST, CERVICAL, AND COLORECTAL CANCER INCIDENCE

Abstract

Objective: To evaluate the geographic variation of breast, cervical, and colorectal cancer incidence in Baltimore City and the neighborhood characteristics associated with the observed variation.

Design and Methods: Using a Maryland Cancer Registry dataset of breast, cervical, and colorectal cancers (N=4,966) diagnosed among Baltimore City female residents from 2000 to 2010, the geographic distribution of incidence was evaluated for each cancer site. The age group of interest was females diagnosed between the ages of 21 to 74 years. Spatial analyses were conducted through the utilization of cholorpleth maps, spatial cluster identification, and local Moran's I. Ordinary least squares regression models were then used to identify neighborhood characteristics associated with the geospatial clusters observed. Cancer stage was also evaluated as a secondary outcome to identify communities in Baltimore City where cancer was more likely to be diagnosed at Stage III or Stage IV.

Results: Each of the cancer sites of interest exhibited geographic variation across Baltimore City. Surprisingly, the observed variability for incidence and stage differed by primary site. Some neighborhoods had a higher absolute burden for one diagnosed cancer but not the others. Specifically, breast cancer had significant low incidence clustering in downtown Baltimore where cervical cancer had a high incidence. Additionally, the significant neighborhood covariates that explained the geographic variation were also different by site. Crime was associated with lower incidence for breast and colorectal

cancer, but not for cervical cancer. It is hypothesized that the younger age distribution in high-crime areas might be driving this association. To a lesser extent, it is possible the type of crime committed is playing a role in a lower life expectancy in high-crime areas. However, this would require additional research. The results from the local Moran's I also provided insight into the location of discordant clusters, which were proximal to communities with a significantly worse or better incidence rate.

Conclusions: Cancer incidence and stage varied geographically even within a single county. Small area estimates are needed when developing health programs to achieve a high enough resolution to detect local patterns of disease. It would also be important to develop place- and outcome-specific interventions given the variability seen by neighborhood in terms of which community-level characteristics were significantly associated with incidence and stage.

Introduction

For several decades, resources and research have been directed towards improving cancer-related outcomes. Lately, these efforts appeared to have yielded noticeable improvements. In recent trends, some cancer sites, such as colorectal and cervical, have experienced annual incidence declines of almost 3% while others, such as breast cancer, have remained stagnant.[1] These developments are a byproduct of improved preventative measures and better screening practices, including those that identify pre-cancerous lesions and polyps for cervical and colorectal cancer, respectively. However, these declines in incidence have not been consistent when stratifying by subgroups within the population. Although African-American females have a lower overall cancer incidence, their rates have remained relatively unchanged from 2000 to 2009 as compared to the 0.3% annual decline seen in white females over the same period.[2] Research has also demonstrated that cancer is diagnosed at a later stage among African-Americans.[3]

The gap observed in cancer incidence and stage distribution varied by cancer site. While the mean incidence of breast cancer was 4% lower in African-Americans as compared to whites from 2005 to 2009, African-Americans had a 34% higher occurrence of cervical cancer.[2,3] It should also be noted that the 34% is a conservative estimate for the difference in cervical cancer burden between African-Americans and whites. These figures likely underestimate cervical cancer incidence in African-Americans by not correcting for hysterectomy status, which is a procedure more prevalent within this racial/ethnic group.[4] The difference also varied by age group as African-American women under the age of 45 years, when compared to white women, had a higher incidence of breast cancer but had comparable cervical cancer incidence among women under 50 years old. At the time of diagnosis, African-American cases were more likely to be diagnosed at either a regional or distant stage for breast, cervical, and colorectal cancer as compared to their white counterparts.[2]

Traditionally, research had primarily centered on the role of individual-level exposures as the cause behind these observed differences in cancer incidence and stage distribution. The focus had been restricted to the utilization of medical and etiological models. More recently, however, there has been a growing interest to evaluate the geography of both cancer and its risk factors.[5–7] The research has consistently shown that cancer outcomes and their known causal exposures exhibit geographic variation that coincide with area-level socioeconomic status and the racial/ethnic composition of neighborhoods.[8–12] It should also be noted that the expansion of the disease paradigm has resulted in directing more attention towards not only the geographic distribution of biological factors, such as environmental toxins, but also of social determinants of health within the scope of cancer.[13–15]

The broadening of the medical model is relevant in two particular ways for cancer research. Firstly, it brings new perspective to the already known epidemiological exposure-outcome associations that have been studied at the individual-level. There is a growing understanding that these well-known associations may vary within different social settings or community context. Secondly, it leads to a better grasp of how distal factors play a role on health outcomes and subsequently provides more upstream avenues on which to intervene.

This chapter aims to evaluate the geographic distribution of breast, cervical, and colorectal cancer incidence among female residents of Baltimore City, Maryland as well as the neighborhood characteristics associated with that distribution. The methodology will incorporate several neighborhood characteristics beyond socioeconomic status and residential segregation, which are the most commonly used area-level measures. The overall purpose of this chapter is to provide exploratory evidence as to the neighborhood context in which cancer incidence tends to occur more frequently and at later stages in Baltimore City.

Methods

Cancer incidence ascertainment

The Maryland Cancer Registry (MCR), described in Chapter 1, served as the data source for the primary outcome, cancer incidence. In order to be eligible for the analysis, cancer cases needed to be a female resident of Baltimore City diagnosed between the ages of 21 to 74 years within the timeframe of 2000 to 2010. Only cases that were classified as having breast, cervical, or colorectal cancer as the primary site were included. Given the spatial nature of the analysis, cases were also required to have a street address that could be geocoded (i.e., assigned latitude and longitude coordinates). The study sample also restricted the cases to only their first occurrence of cancer. For example, if the same individual developed cancer multiple times in the same primary tumor site, only the first diagnosis was counted. For instances where the same individual had a primary diagnosis in multiple sites of interest (e.g., cervical and breast cancer), each tumor was counted separately and evaluated in the corresponding cancer site-specific model. Information on cancer stage was also obtained and examined as a secondary outcome.

Neighborhood characteristics

Also discussed in Chapter 1, the Vital Signs report was utilized for the independent neighborhood-level variables that described the quality of life within each Community Statistical Area (CSA).[16] This publicly available dataset characterized each community through 66 indicators that spanned across 7 domains. Of the 66 indicators available, the list was reduced to those related to social context, social relationships, and physical context, as described in the adapted Warnecke model from Chapter 1. A few methods were considered in order to reduce the number of indicators down to a list with the most pertinent community characteristics. One approach would have been to include all 66 indicators into the initial exploration of the univariate relationships between each CSA indicator and the relevant cancer outcome. From there, model-based techniques would have been used to identify the most parsimonious and best fitting model. However, the integration of such an exhaustive list and the evaluation of each neighborhood indicator could have resulted in seeing a significant association by chance.

As a result, the Warnecke model was first used to identify CSA characteristics that were explicitly identified in the model as well as the Vital Signs report. This overlap included indicators, such as racial/ethnic integration, employment, and social/economic gradient.

After selecting the selecting the characteristics that were clearly noted in the framework, the remaining characteristics in the Vital Signs report were assessed for potential proxies that could serve as correlates for other factors that were referenced in the Warnecke Model. These proxies were validated with literature that demonstrated a correlation between the available neighborhood variable and the physical/social context mentioned in the paradigm. This resulted in the inclusion of tree coverage, which has been shown to be a marker of social cohesion, and vacant housing, which has an inverse relationship with neighborhood stability.[17,18]

For the reduced list, each indicator's availability by year along with its corresponding domain, definition, and units can be found in **Appendix 2A**. Trend line plots were created for this subset to determine whether they remained stable over time within a CSA. If an indicator displayed consistency in either its absolute figure or ranking across the CSAs, the data were averaged across the available years. Additionally, a correlation matrix was created to evaluate whether the narrowed list exhibited repetitiveness within a domain as well as to understand the correlation of indicators across domains. While indicators within a domain can not be highly concurrent, correlation is allowed for indicators in different domains. This method minimized redundancy and ensured that indicators captured a different aspect of a domain. The indicators in the reduced list were explored for associations with cancer incidence as well as cancer stage using the methodology described below.

Statistical analyses

A series of descriptive statistics were conducted across the aggregated cancers as well as stratified by site on the cohort of female residents with a cancer diagnosis using STATA 12.1.[19] For individual-level characteristics provided by the MCR, mean age at diagnosis, tumor grade, race, and treatment distributions were evaluated. At the CSA-level, the total number of cancer cases within each geographic boundary was determined. Choropleth maps were created to shade CSAs for each cancer site by quintiles in terms of cancer incidence and cancer stage. These thematic maps of epidemiologic metrics provided an exploratory overview of Baltimore City neighborhoods that were in the 80th percentile for incidence and stage distribution. The statistical significance of any spatial cluster was analyzed through the Getis-Ord Gi* statistic, which identified hot spots. By evaluating the local sum of a CSA and its neighbors relative to the total sum in all of Baltimore City, this local spatial method measured pockets of spatial association that may have otherwise been obscured by global statistics (e.g., obtaining a cancer incidence rate only at the county-level).[20]

Once the areas that produced more or less than expected events were identified using the Getis-Ord Gi* statistic, a global ordinary least squares (OLS) regression model was conducted for each cancer site to identify potential independent variables that explained the geographic distribution of these clusters. The OLS model is the starting point of all spatial regression analyses and provides a single regression fit that begins to explain what neighborhood characteristics drive cancer to occur where it does. This model produces a linear fit that yields the smallest residuals. Finally, a local indicators of spatial association

(LISA) function using Local Moran's I was conducted to analyze: 1) the extent to which each neighborhood influenced the global magnitude of cancer incidence as well as cancer stage across all of Baltimore City; and 2) the extent to which each neighborhood was surrounded by other neighborhoods with similar values.[21,22] This method not only illustrated spatial clusters but also identified discordant clusters, such as CSAs with high cancer incidence surrounded by others with low incidence. The examination of any neighboring clusters that were dissimilar is particularly relevant since information on the potential drivers of those differences could be leveraged into the development of interventions. All of the described spatial analyses was conducted using ArcGIS 10.3 software.[23]

The Johns Hopkins School of Public Health's Institutional Review Board and the Maryland Department of Health and Mental Hygiene's Institutional Review Board determined this study to be exempt research.

Results

Neighborhood characteristics

Of the original 66 indicators available through the Vital Signs report, the list was reduced to 15 indicators that best described the social relationships as well as the social and physical context of environments as defined in the Warnecke model. For many of these CSA characteristics, multiple years of data were available. Trend lines were used to evaluate whether the characteristics noticeably changed over time and would require including year-specific data in the models. The trend lines demonstrated that the indicators remained relatively stable over the 2000 to 2010 study period. The trend line plots for all 15 indicators are available in **Appendix 2B** for reference. Even when there were slight increases or decreases in the covariates over time, the ranking across CSAs remained mostly constant. As a result, the average was taken across all available years of data and utilized as a single summary measure for each CSA. The average values of all 15 indicators utilized in the subsequent spatial analysis for each of the 55 CSAs can be found in **Appendix 2C**.

Of the neighborhood-level covariates provided by the Vital Signs report, the 15 indicators included were classified by domain and provided within a correlation matrix (**Table 2.1**). The highest correlations within a domain were seen between single female-headed households and percent African-American residents (0.79) as well as between single female-headed households and percent of households with less than a $25,000 income (0.80). Since the indicator of single female-headed households was highly correlated with two indicators pertaining to the same domain, it was given higher precedence for being included in any of the subsequent regression models because it appeared to track closely with both the racial/ethnic and socioeconomic gradient of a CSA.

Cancer incidence and population characteristics

The study cohort used in the analysis consisted of 4,966 total cases across the three cancer sites of interest: breast (n= 3,466), cervical (n= 380), and colorectal (n=1,120) cancer. There were 181 cases dropped from the initial sample provided by the MCR due to having an address that could not be geocoded. For most of these exclusions, the primary reasons consisted of having no street address available (e.g., only a P.O. Box was listed) or residing in a county outside of Baltimore City.

As seen in the descriptive characteristics (**Table 2.2**), the entire cohort was diagnosed at a mean age of 56.3 (SD= 11.2) years. However, the mean age at diagnosis significantly differed across cancer sites. Cervical cancer cases were diagnosed, on average, at a younger age of 50.0 (SD= 12.7) years while colorectal cancer cases were the oldest of the three sites at 59.9 (SD= 10.1) years. When comparing across cancer sites, breast cancer had a higher proportion of cases classified as Grade III or higher but had a lower proportion of deaths. Significant differences were observed for cancer stage and vital status across sites while racial/ethnic groups were similarly distributed.

The utilization of treatment therapies varied across cancer types as well. It should be noted that, given the data source, therapy type had the highest proportion of missing or unknown values compared to the other individual-level data provided by the MCR. Across the cancer sites, chemotherapy and hormone therapy usage consistently had at least 20% missing values for these fields. For example, in cervical cancer, it was not known for 22% of the cases whether they received chemotherapy. The cancer registry does not regularly follow-up nor verify treatment patterns among diagnosed cases. This has etiological relevance at the individual-level since the receipt of timely and high-quality treatment can be influential in the prognosis of cancer. The community-level context in which therapy occurs can potentially modify the individual-level association between treatment and progression of disease. However, there is limited information on how social context can impact treatment utilization and/or treatment response.

In terms of cancer incidence by CSA, the number of cases ranged from 27 to 191 diagnosed residents in any single geographic unit (**Table 2.3)**. Cedonia/Frankford yielded the most cancer cases while Dickeyville/Franklintown had the fewest. A table was reproduced from Chapter 1 to assist in identifying CSAs by name (**Table 2.4**) in the subsequent maps. The choropleth maps demonstrated that the geographic distribution of cancer incidence per 1,000 female residents aged 21 to 74 years appeared to vary by cancer site (**Figures 2.1a-d**). Breast cancer incidence appeared to aggregate in the northeast and northwest areas of Baltimore City while cervical cancer mostly occurred in the southeast and southwest CSAs. The highest quintiles of colorectal cancer incidence were located in similar areas as breast cancer with additional high incidence in the southwestern CSA of Westport/Mount Winnans/Lakeland. These thematic maps of cancer incidence begin to provide an early indication that epidemiologists should take a more local approach when evaluating patterns and distributions of disease. There is often too much reliance on national surveys or broad city-level measures. Cities are rarely, if ever, equally homogenous in terms of disease burden or risk factors. When small area estimates are not taken into consideration, public health professionals run the risk of applying approaches that are costly and ineffective because they were implemented in the wrong neighborhood.

These choropleth maps were also utilized to portray cancer incidence per 1,000 female residents aged 50 to 74 years and are displayed in **Appendix 2D**. The change in the denominator's age group did not appear to affect the distribution of cancer incidence. Of the cancer cases provided by the MCR, 1,412 cases had an age of less than 50 years at the time of diagnosis. Although the older age group of 50 to 74 years was a higher-risk group, the main analysis utilized the broader age range of 21 to 74 years since the at-risk population included cancer cases within that age range. Additionally, cancer screening across the three sites of interest targets an age range of 21 to 74 years due to the screening guidelines for cervical cancer. While the main analysis for this chapter utilized the 21 to 74 years age range, the analysis was also conducted with a more restricted age group of 50 to 74 years. The results for this additional geospatial analysis can be found in **Appendix 2D** for informational purposes. Overall, the restriction on age did not drastically change the results observed in the spatial analyses. The location of spatial clusters remained consistent for each of the geostatistical methods utilized (i.e., choropleth maps, hot spot analysis, and local Moran's I).

In order to make a more fair comparison of geographic distributions across cancer sites, Z-scores of cancer incidence by site were also mapped using the 21 to 74 years age group. Based on the Z-score, the density of breast cancer increased throughout Baltimore City. The northeast and northwest CSAs continued to have the highest burden of incidence. For cervical cancer, the conversion to Z-scores removed some of the density in central Baltimore City while colorectal cancer saw its highest burden areas remain relatively unchanged. The thematic maps that used Z-scores are available for review in **Appendix 2E**.

Spatial analysis of cancer incidence

The results of the hot spot analysis assessed the statistical significance of the patterns observed in the choropleth maps. Each hot spot analysis used the cancer incidence per 1,000 female residents aged 21 to 74 years as the value to compare to neighboring CSAs. As expected, the hotspots across the aggregated cancer incidence map resembled that of breast cancer (**Figures 2.2a-d**). Since the cohort had significantly more breast cancer cases than either colorectal or cervical cancer, the distribution of the overall incidence was driven by breast cancer. When comparing across the three cancer sites, there were several differences in terms of where the significant hot and cold spots were located within Baltimore City. The northern area of Baltimore City had a significant aggregation of breast cancer that was not seen in either cervical or colorectal cancer. In fact, the two cold spots yielded by the cervical cancer analysis were located in the area of high breast cancer incidence. Conversely, cervical cancer had a significant clustering of cases in Canton, which was a neighborhood with significantly lower breast and colorectal cancer incidence.

Upon testing the significance of the observed spatial pockets of incidence, it becomes more evident through the varying distributions that risk factors appear to operate differently within CSAs. A heat map was provided (**Table 2.5**) to visually demonstrate that a CSA falling in the highest quintile for incidence in one cancer site did not necessarily fall in the highest quintile for another cancer site. For example, Harbor East/Little Italy was in the top 80th percentile for cervical cancer incidence while falling to the bottom 20% for breast cancer. Of the 55 CSAs, only 10 neighborhoods had each of the three cancer sites appear in the same quintile. Three of these communities, Belair-Edson, Midway/Coldstream, and Southern Park Heights, had all three cancer incidence rates in the highest quintile. This result strengthens the argument that local drivers may go unnoticed if there is too much aggregation, such as collapsing across all primary tumor sites with the assumption that the same neighborhoods will be appear as high-risk clusters for all cancer types.

In order to explain the geographic location of hot and cold spots for each cancer site, an OLS regression model was carried out using each of the neighborhood-level characteristics listed in the previously described correlation matrix (**Table 2.1**). These 15 covariates were selected from the 66 indicators made available by the Vital Signs report because of their alignment with the Warnecke framework for cancer disparities. They are either a direct measurement of, or a proxy for, the intermediate factors of social context, social relationships, and physical context, which lead to disparate cancer outcomes.

The coefficients yielded by the models can be found in **Table 2.6**. When stratifying by cancer site, the significant associations with each of the CSA characteristics vary. For all cancers, an increased proportion of African Americans residents within a CSA as well as a decrease in number of businesses were associated with higher cancer incidence. As was briefly discussed in Chapter 1, black non-Hispanic females have a lower breast cancer incidence as compared to white non-Hispanics at the individual-level.[24] This analysis indicated that every one percent increase in African-American residents resulted in breast cancer incidence increasing by 0.059 times per 1,000 female residents. While these associations may appear to be contradictory, it should be noted that given the use of the community-level variables in this analysis, cross-level inferences should not be made in order to avoid ecologic fallacy. Although there is a significant association between an increased proportion of African-American residents and increased breast cancer incidence, this does not necessarily mean that the African-American residents are the ones developing breast cancer.

Crime was shown to have an inverse relationship indicating that CSAs with less crime had a higher cancer burden, which was not an expected relationship. One possible explanation is that high-crime areas tend to have a younger population. As a result, cancer incidence would appear to be lower given that older individuals make up a small portion of the population. This significant association was observed in both breast and colorectal cancer. This relationship may not have been observed for cervical cancer due to the younger average age at diagnosis as compared to the other sites. Since the crime indicator encompasses all Part 1 offenses (criminal homicide, forcible rape, theft, assault, and arson), additional research is required to determine if the type of crime being committed is lowering the life expectancy in CSAs with prevalent crime. To a lesser extent, it is possible the type of crime committed is playing a role in a lower life expectancy in high-crime areas. However, this would require additional research.

Surprisingly, a higher proportion of voters, which is often a proxy for the sustainability and social capital of a neighborhood, appeared to be significantly associated with a greater breast cancer burden. This can likely be explained through voter participation occurring in more urbanized settings where voting facilities are more easily accessible. These settings tend to be densely populated and thus resulting in more cancer cases. Overall, many of the significant associations observed in the unadjusted models were expected given the potential pathways these intermediate factors could affect cancer incidence. For cervical cancer, neighborhood characteristics, such as a high teen birth rate, reflected the behavioral risk factors of unprotected sex, which is a predictor of cervical cancer.

Each of these models was evaluated through several OLS diagnostic tools to determine fit as well as to avoid bias. None of the unadjusted spatial models were statistically significant for the Jarque-Bera statistic, which would have indicated that the residuals were not normally distributed. The significant covariates for each cancer site were integrated into their respective adjusted models (**Table 2.6**). While not significant in the unadjusted model, the number of females 50 to 74 years that lived within a CSA was included in the adjusted model due to the known association between age and cancer risk. Most of the neighborhood covariates retained their statistical significance in the fully adjusted model, except for cervical cancer. The integration of the covariates in that case resulted in null associations. The final models were evaluated through the R-squared statistic, which demonstrated the proportion of the variability in the outcome that was explained by the model. For breast cancer, the model fitting percent of females 50 to 74 years, percent of African-American residents, crime rate, number of businesses, percent of population that voted, and tree coverage explained over one-third of the geographic distribution observed in the hot spot maps.

These final models were also mapped out to provide a visualization of how the model adjustment affected the variability of the outcome. As seen in **Figures 2.3a-d**, the residuals of the OLS models were mapped out by standard deviation. Areas highlighted in red were areas that had more than expected cancer cases as compared to what was predicted by the final model. Conversely, areas in blue had a lower than expected cancer incidence. Similar to non-spatial statistics, the residuals of a model should be randomly distributed. If the breast cancer hot spot analysis (**Figure 2.2b**) was to be compared to the final OLS model for breast cancer (**Figure 2.3b**), the residuals of the model appear to be more randomly distributed across Baltimore City than the high incidence cluster previously seen in the northern area. This randomness serves as an indicator that the model produced randomly distributed error, which is preferred when evaluating a model's fit.

The final spatial tool used to evaluate cancer incidence was the local Moran's I. The strength of this method is that it demonstrates how the small area estimates contribute to the global measure while also highlighting concordant and discordant clusters. As seen in **Figures 2.4a-d**, high-low and low-high clusters were observed. In the case of cervical cancer, Southern Park Heights was shown to be a high-low cluster, indicating that this particular CSA was unusual in its location given that it had higher cervical cancer incidence while being an immediate neighbor to an area with low incidence. These discordant clusters have the potential to serve as a starting point for future research by delving into the driving factors that result in two proximal CSAs having such different risk profiles.

Spatial analysis of cancer stage

For cancer stage by CSA, the number of stage 4 cases ranged from 0 to 20 (**Table 2.3**). Cedonia/Frankford (n = 20) and Dickeyville/Franklintown (n = 0) yielded the most and fewest stage 4 cases, respectively, which was expected given how many residents were diagnosed with cancer in each of them. Each CSA was described according to the proportion of cancer cases that were either stage 3 or 4. Similar to cancer incidence, the choropleth maps demonstrated that the geographic distribution of stage 3 and 4 cancers varied by cancer site (**Figures 2.5a-d**). The distribution differed even within cancer sites when comparing cancer incidence to stage. In the case of breast cancer, a higher proportion of cases in the central area of Baltimore City were diagnosed at a later cancer stage. This was unlike the distribution of cancer incidence overall, which had appeared to become more concentrated in the northeast and northwest portions of the county. The initial differences observed within the same cancer site for incidence and stage once again indicate that different local factors may have varying associations for different cancer-specific outcomes. Upon using the Z-scores, the distribution of cancer stage across the sites became more similar as seen in **Appendix 2F**.

Unexpectedly, there was a cancer stage hot spot in the downtown Baltimore City area for breast cancer, which had been a cold spot for breast cancer incidence (**Figures 2.6a-d)**. This was unlike colorectal cancer, which had its hot spots for cancer incidence and cancer stage for the most part coincide with each other. A heat map is again provided (**Table 2.**7) to draw attention to noteworthy situations, such as those that occur within Northwood where a higher burden of late stage is seen for cervical cancer but not for breast or colorectal. In total, only 7 CSAs had cancer stage appear in the same quintile for each of the three cancer sites. These differences begin to demonstrate that interventions need to be both geographically tailored and well-developed in terms of the outcomes used to ascertain effectiveness.

OLS models were used to explain the geographic variation observed in stage (**Table 2.8**). Cervical cancer had no significant associations with any of the 15 indicators. Crime was no longer a statistically significant indicator for any of the cancer sites while domestic violence had significant associations with the overall distribution of cancer stage as well as with breast cancer's distribution. Its direction indicated that areas with a high call volume for domestic violence were also areas that had a higher proportion of cases diagnosed at either stage 3 or 4. The mapping results of the OLS models showed some clustering of the residuals, which indicates that the model was missing key variables that would have explained more of the geographic variation observed (**Figures 2.6a-c**).

The local Moran's I yielded a low-low cluster that overlapped for both breast and colorectal cancer in North Baltimore/Guilford/Homeland (**Figures 2.7a-d**). This neighborhood would appear to have some characteristics conducive to lower stage cancers, especially considering that it was one of the CSAs that yielded the most cases. As previously mentioned, these results prompt the need for further investigation in understanding the mechanism that protects some neighborhoods and not others. For this CSA, one potential explanation for the lower stage cancers despite yielding the most cases is there might be a neighborhood infrastructure that allows for more access to screening. Prompt screening and adherence to guidelines increases the likelihood of detecting cancer in its earlier stages when it is more treatable.

The entire cancer stage geospatial analysis was replicated using the age restriction of females aged 50 to 74 years at the time of diagnosis and is available in **Appendix 2G**. In this additional analysis, the choropleth map had fewer areas of high burden as more of the neighborhoods fell within the 50th percentile. While the hot spot analysis did overlap with the results of the broader age range, the OLS models yielded different significant associations. As mentioned in the OLS results, cervical cancer was not significantly associated with any of the community-level covariates when implementing the age range of 21 to 74 years. However, when restricting the group to females aged 50 to 74 years, it was colorectal cancer that failed to have any significant associations. This demonstrates the need for a clearly defined target population when developing programs in order to identify the relevant community characteristics to intervene on specific to the group of interest.

Conclusion

Overall, the three primary findings in this chapter were: 1) the clear existence of geographic variation in cancer burden within Baltimore City in terms of incidence and stage; 2) the noticeable difference in this observed geographic distribution for breast, cervical, and colorectal cancer; and 3) the difference in statistically significant neighborhood characteristics that explained some of the geographic variation across the three sites. In the evaluation of the exploratory thematic maps and hot spot analysis, different areas across Baltimore City appeared to have more cancer outcomes than others, which aligned with the study's a priori hypothesis. However, the observed difference in the location of spatial clusters when comparing across cancer sites was not expected. While each primary cancer was anticipated to have distinct nuances and possess some hot spots not observed in other cancers, the instances in which some neighborhoods were a cold spot for one cancer but a hot spot for another was surprising. Even when evaluating within a cancer site and comparing across different cancer outcomes, such as cancer incidence versus cancer stage, there were differences observed. This would indicate that developing geographically tailored interventions to address cancer would not only need to be site-specific but also outcome-specific. This finding demonstrates that there may have likely been past cancer-related interventions that, although might have appeared ineffective, were actually misallocated to incorrect populations and geographic locations that did not fit the targeted risk profile.

Additionally, the finding, which demonstrated different associations with neighborhood characteristics across cancer sites, also warrants careful consideration when attempting to intervene through more upstream avenues. The assumption might have previously been made that addressing a particular social determinant or improving a specific contextual neighborhood characteristic would result in downstream improvements for all health outcomes. For example, previous literature has shown that cancer incidence is differentially distributed across a geographic region.[25,26] The findings in this dissertation are consistent with the geospatial work that has been done up until now in that regard. Unfortunately, the previously available research relied primarily on Census data, which is collected every decade, to describe the regions and also used administrative boundaries, such as zip codes and county lines, to define place, which results in very heterogeneous geographic units within the analysis. The unique community measures in this dissertation provide a new perspective on the relationship between social determinants at the neighborhood level and subsequent health outcomes. This chapter provides early evidence that neighborhood factors might affect exposure-outcome associations in different ways. Without this understanding, this would again result in null findings for interventions if the appropriate benchmarks or outcomes used to measure effectiveness were not selected.

This chapter has several strengths and limitations. One limitation of this study is that neighborhood context is limited only to the residential environment, as defined by the address at the time of the cancer diagnosis, of the study population. The concept of social determinants is an all-encompassing one that evaluates how health "is impacted by where and how we live, learn, work, and play."[27] Although place of residence makes up a large component of the environment individuals interact with, it fails to capture the additional resources or disadvantages an individual may be exposed to through the areas in which they work or socialize. As a result, the CSA designation each cancer case within the study population was categorized into may not fully represent their neighborhood context.

Another limitation related to the residential address data is that the MCR did not have information on how long female cancer cases had resided at their reported address prior to their diagnosis. Since cancer is recognized as having a long latency period, an incident cancer case that only recently moved to her current residence in Baltimore City may not have been exposed long enough to her neighborhood to be a true representation of that area's association with cancer outcomes. The same can also be said for individuals that have lived in Baltimore City for several decades but may have left just prior to receiving a cancer diagnosis. It is unlikely that one of these two scenarios occurred more frequently than the other as populations naturally fluctuate over time. Upon assessing the data available through the U.S. Census Bureau, over 80% of female residents between the ages of 21 to 74 years reported having the same residence as the year before. This estimate remained relatively stable throughout the study period of 2000 to 2010.[28] Most of the mobility came from the younger age groups as this figure ranged from 65% to 78% among women between the ages of 21 to 33 years. For the most part, the population was relatively stable, despite the limited information on those who had left Baltimore City and those that had only recently arrived prior to their diagnosis date.

Another aspect of this study that might be viewed as a limitation is the absence of individual-level risk factors, such as age and race, in the OLS regression analyses.

Without these variables, the models are limited in determining how much of the association seen between neighborhood characteristics and cancer incidence is due to individual-level factors versus how much is due to the independent effect of the CSA. While in certain scenarios, particularly those with etiological objectives, it is important to distinguish between individual versus environmental factors. The intricate nature of this question was beyond the scope and overarching research question of this study. The focus was to validly assess whether the social and physical environments of Baltimore City neighborhoods were indicators for cancer incidence and stage. This information is particularly relevant for future interventions that want to allocate resources and services to areas that appear to have a higher absolute public health burden (e.g., having more residents diagnosed with late stage cancer). This process may be done more efficiently if neighborhood characteristics can be used as markers for susceptible communities as opposed to collecting individual-level data to obtain a more refined look at the composition of residents. This study takes a population-based approach in addressing cancer disparities by taking the broader perspective of neighborhood context.

Additionally, the reliance on neighborhood-level indicators rather than individual-level data is a realistic scenario encountered by health departments as well as local organizations in the development and operationalization of programs to address community health. As was briefly discussed in the results section, cross-level inferences should be avoided given the use of only community-level characteristics. For example, the association between having a high proportion of African-American residents within a CSA and having higher cancer incidence within that same CSA does not necessarily mean that the African-American residents are the ones being diagnosed with cancer at a higher rate.

Despite these limitations, this study has a number of strengths. Firstly, it improves upon prior characterizations of neighborhood context by integrating community-level measures that have been collected on Baltimore City neighborhoods for over a decade by the Baltimore Neighborhood Indicators Alliance.[16] These unique variables came about as the Baltimore City Health Department compiled and disseminated its neighborhood profile assessments. Previous studies have relied mostly on data collected by the U.S. Census Bureau, which may not have extensively measured the social and physical environments of communities. While others may have attempted to account for these classifications in their analyses, few have utilized community covariates that have the same breadth and depth as those tracked in Baltimore City. The use of these additional neighborhood data painted a more complete picture of where Baltimore City residents live while also providing potential suggestions as to neighborhood characteristics that should be collected more routinely in other cities.

Additionally, through the use of the Baltimore Neighborhood Indicators Alliance, the study was able to utilize the CSA as its geographic unit of analysis. This was beneficial in two ways. One benefit was that the ability to focus on neighborhoods within a specific county allowed for the calculation of small area estimates. These estimates improved statistical precision while also maintaining geographic resolution. The results provided small-scale patterns within neighborhoods but also large-scale trends across the whole county. Small area estimates increase the likelihood of homogeneity across the neighborhood characteristics within the geographic boundaries and thus increase the validity of the associations. There have been other studies on geographic variation across broader regions using larger datasets that may have overlooked the local patterns of disease, especially in the absence of some of the unique community data utilized in this study.[29,30] The second benefit consists of the cultural relevance of the geographic boundaries utilized. Residents are often unaware of the census tracts they live in, which is another small-scale administrative unit often used in geospatial analyses. Through the Baltimore Neighborhood Indicators Alliance, Baltimore City residents have grown accustomed to hearing about the quality of life within these boundaries due to the measures that have been collected over the years. As a result, CSAs might be easier to identify with and subsequently intervene on since there is a better grasp of the population's composition and context.

This study took a transdisciplinary approach, which is often underutilized in studies that focus mostly on clinical associations. By evaluating cancer incidence and cancer stage within a neighborhood context, the study cut across diverse disciplines. By taking into account neighborhood residence, the findings help shed light on the complexity of cancer disparities and provide additional insight as to the research areas that should join in collaborative efforts to address inequities in health outcomes among subgroups. Based on the results, there is compelling evidence to pursue further research on the association of neighborhood factors with the geographic distribution of cancer incidence and stage beyond just Baltimore City. The results of the local Moran's I would be particularly useful in focusing on discordant neighboring CSAs. The findings make the case that there is an opportunity to create effective geographically tailored cancer services, which has been accomplished in other communities.[31–33]

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015;65(1):5–29.

2. Jemal A, Simard EP, Dorell C, et al. Annual Report to the Nation on the Status of Cancer, 1975–2009, Featuring the Burden and Trends in Human Papillomavirus (HPV)–Associated Cancers and HPV Vaccination Coverage Levels. JNCI J Natl Cancer Inst 2013;:djs491.

3. Howlader N, Noone AM, Krapcho M, et al., eds. SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations). [Internet]. Bethesda, MD: National Cancer Institute; 2011. Available from: http://seer.cancer.gov/csr/1975_2009_pops09/

4. Rositch AF, Nowak RG, Gravitt PE. Increased age and race-specific incidence of cervical cancer after correction for hysterectomy prevalence in the United States from 2000 to 2009. Cancer 2014;120(13):2032–8.

5. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. Am J Epidemiol 2002;156(5):471–82.

6. Singh GK, Miller BA, Hankey BF, Edwards BK. Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975–1999. Bethesda, MD: National Cancer Institute; 2003.

7. Baker null, Hoel null, Mohr null, Lipsitz null, Lackland null. Racial, age, and rural/urban disparity in cervical cancer incidence. Ann Epidemiol 2000;10(7):466–7.

8. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. Br J Cancer 2002;87(6):585–90.

9. Coughlin SS, Leadbetter S, Richards T, Sabatino SA. Contextual analysis of breast and cervical cancer screening and factors associated with health care access among United States women, 2002. Soc Sci Med 2008;66(2):260–75.

10. Tatalovich Z, Zhu L, Rolin A, Lewis DR, Harlan LC, Winn DM. Geographic disparities in late stage breast cancer incidence: results from eight states in the United States. Int J Health Geogr 2015;14(1):31.

11. Klassen AC, Pankiewicz A, Hsieh S, Ward A, Curriero FC. The association of area-level social class and tobacco use with adverse breast cancer characteristics among white and black women: evidence from Maryland, 1992-2003. Int J Health Geogr 2015;14:13.

12. Crabbe JCF, Gregorio DI, Samociuk H, Swede H. Secular trends, race, and geographic disparity of early-stage breast cancer incidence: 25 years of surveillance in Connecticut. Am J Public Health 2015;105 Suppl 3:e64–70.

13. Gerend MA, Pai M. Social Determinants of Black-White Disparities in Breast Cancer Mortality: A Review. Cancer Epidemiol Biomarkers Prev 2008;17(11):2913–23.

14. Hiatt RA, Breen N. The Social Determinants of Cancer. American Journal of Preventive Medicine 2008;35(2):S141–S150.

15. Williams DR, Jackson PB. Social Sources Of Racial Disparities In Health. Health Aff 2005;24(2):325–34.

16. Baltimore Neighborhood Indicators Alliance–Jacob France Institute. Vital Signs 10: Measuring Baltimore's Continued Progress Towards Strong Neighborhoods and a Thriving City. Baltimore, MD: University of Baltimore; 2012.

17. Schwarz K, Fragkias M, Boone CG, et al. Trees Grow on Money: Urban Tree Canopy Cover and Environmental Justice. PLOS ONE 2015;10(4):e0122051.

18. What Do the Neighbors Think? Assessing the Community Impact of Neighborhood Stabilization Efforts - Boston Fed [Internet]. [cited 2016 Mar 31];Available from: https://www.bostonfed.org/commdev/necd/2012/issue1/what-do-the-neighbors-think.htm

19. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.

20. Ord JK, Getis A. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 1995;27(4):286–306.

21. Anselin L. Local Indicators of Spatial Association—LISA. Geographical Analysis 1995;27(2):93–115.

22. Moraga P, Montes F. Detection of spatial disease clusters with LISA functions. Stat Med 2011;30(10):1057–71.

23. ESRI. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute; 2011.

24. National Institutes of Health. State Cancer Profiles: Quick Profiles (Maryland) [Internet]. National Cancer Institute; Available from: http://statecancerprofiles.cancer.gov/quick-profiles

25. Rushton G, Peleg I, Banerjee A, Smith G, West M. Analyzing geographic patterns of disease incidence: rates of late-stage colorectal cancer in Iowa. J Med Syst 2004;28(3):223–36.

26. McElroy JA, Remington PL, Gangnon RE, Hariharan L, Andersen LD. Identifying geographic disparities in the early detection of breast cancer using a geographic information system. Prev Chronic Dis 2006;3(1):A10.

27. Riley RW. Health Starts Where We Learn [Internet]. Robert Wood Johnson Foundation. [cited 2015 Nov 11];Available from: http://www.rwjf.org/en/library/research/2010/10/health-starts-where-we-learn.html

28. U.S. Census Bureau. Geographic Mobility in the Past Year by Age for Current Residence in the Unite States: 2006-2010 American Community Survey 5-Year Estimates. U.S. Census Bureau; 2010.

29. Flores YN, Davidson PL, Nakazono TT, Carreon DC, Mojica CM, Bastani R. Neighborhood socio-economic disadvantage and race/ethnicity as predictors of breast cancer stage at diagnosis. BMC Public Health 2013;13:1061.

30. Feinglass J, Rydzewski N, Yang A. The socioeconomic gradient in all-cause mortality for women with breast cancer: findings from the 1998 to 2006 National Cancer Data Base with follow-up through 2011. Ann Epidemiol 2015;25(8):549–55.

31. Alcaraz KI, Kreuter MW, Bryan RP. Use of GIS to identify optimal settings for cancer prevention and control in African American communities. Prev Med 2009;49(1):54–7.

32. Brantley-Sieders DM, Fan K-H, Deming-Halverson SL, Shyr Y, Cook RS. Local breast cancer spatial patterning: a tool for community health resource allocation to address local disparities in breast cancer mortality. PLoS ONE 2012;7(9):e45238.

33. King SJ, Livingston PM, Turner L, et al. Mapping the information resources available to patients with colorectal cancer. Eur J Cancer Care (Engl) 2010;19(4):492–500.

Tables and Figures:

Table 2.1: Correlation matrix by domain of CSA characteristics relevant to social context, social relationships, and physical context

		Census						Housing & Community Development		Crime		Children & Family Health	Workforce & Economic Development		Sustainability			
		Females 21-49	**Females 50-74**	**Percent African American**	**Racial Diversity Index**	**Household Income <25K**	**Female Headed**	**Vacants**	**Housing violations**	**Crime**	**Domestic Violence**	**Teen Births**	**Employed**	**Businesses**	**Voted**	**Dirty Streets**	**Tree Coverage**	**Neighborhood Associations**
Census	**Females 21-49**																	
	Females 50-74	0.88																
	Percent African American	-0.09	-0.03															
	Racial Diversity Index	0.12	0.05	-0.58														
	Household Income <25K	-0.05	-0.05	0.59	-0.25													
	Female Headed	-0.10	-0.10	0.79	-0.41	0.80												
Housing & Community Development	**Vacants**	-0.07	0.16	0.44	-0.41	0.66	0.64											
	Housing Violations	-0.11	0.11	0.60	-0.41	0.77	0.71	0.76										
Crime	**Crime**	-0.07	-0.15	-0.11	0.28	0.20	0.17	0.12	0.20									
	Domestic Violence	-0.11	-0.11	0.58	-0.18	0.73	0.83	0.48	0.54	0.31								
Children & Family Health	**Teen Births**	-0.06	-0.08	0.45	-0.19	0.55	0.65	0.46	0.45	0.14	0.81							
Workforce & Economic Development	**Employed**	0.01	-0.02	-0.59	0.39	-0.79	-0.81	-0.74	-0.74	-0.15	-0.69	-0.52						
	Businesses	0.12	0.04	-0.29	0.40	0.05	-0.04	-0.02	0.05	0.91	0.06	-0.11	-0.01					
Sustainability	**Voted**	-0.18	-0.10	-0.19	-0.11	-0.64	-0.57	-0.49	-0.49	-0.41	-0.69	-0.63	0.57	-0.29				
	Dirty Streets	0.08	0.06	0.32	-0.13	0.34	0.35	0.60	0.41	0.08	0.45	0.48	-0.44	-0.08	-0.41			
	Tree Coverage	-0.10	-0.02	0.02	-0.10	-0.35	-0.26	-0.41	-0.28	-0.42	-0.39	-0.46	0.30	-0.28	0.58	-0.51		
	Neighborhood Associations	0.30	0.32	0.32	-0.39	0.33	0.28	0.40	0.44	-0.03	0.10	0.18	-0.44	-0.01	-0.13	0.16	-0.07	

Table 2.2: Descriptive characteristics of female cancer cases in Baltimore City, MD (2000 to 2010)

	Breast (n= 3,466)	Cervical (n= 380)	Colorectal (n= 1,120)	All sites (N= 4, 966)	p-value
Mean age at diagnosis-years (SD)	55.8 (11.0)	50.0 (12.7)	59.9 (10.1)	56.3 (11.2)	<0.001*
Cancer stage- % (n)					
Stage 0	1.5 (52)	2.1 (8)	1.3 (15)	1.5 (75)	<0.001*
Stage I	46.3 (1,606)	39.5 (150)	31.3 (350)	42.4 (2,106)	
Stage II	17.6 (609)	15.0 (57)	13.5 (151)	16.5 (817)	
Stage III	21.8 (754)	20.5 (78)	20.4 (228)	21.4 (1,060)	
Stage IV	6.0 (208)	9.7 (37)	22.1 (248)	9.9 (493)	
Tumor grade- % (n)					
Grade I	13.4 (464)	9.5 (36)	8.5 (95)	12.0 (595)	<0.001*
Grade II	32.5 (1,125)	30.0 (114)	56.1 (629)	37.6 (1,868)	
Grade III	41.7 (1,446)	25.8 (98)	16.4 (183)	34.8 (1,727)	
Grade IV	0.6 (21)	1.3 (5)	1.0 (11)	0.7 (37)	
Vital Status- % (n)					
Deceased	29.8 (1,033)	53.9 (205)	50.3 (564)	36.3 (1,803)	<0.001*
Alive	70.2 (2,433)	46.1 (175)	49.7 (557)	63.7 (3,167)	
Race					
White Non-Hispanic	31.0 (1,075)	27.6 (105)	27.5 (308)	30.0 (1,488)	0.064
Black Non-Hispanic	58.5 (2,027)	61.8 (235)	63.5 (711)	59.9 (2,973)	
Other	1.9 (65)	1.8 (7)	1.4 (16)	1.8 (88)	
Chemotherapy- % (n)					
Yes	43.7 (1,515)	39.0 (148)	37.0 (414)	41.8 (2,077)	0.001*
No	35.6 (1,235)	39.0 (148)	39.2 (439)	36.7 (1,822)	
Hormone therapy- % (n)					
Yes	25.9 (899)	<1.6 (<6)	0.0 (0)	-[a]	<0.001*
No	50.6 (1,754)	80.0 (304)	81.3 (911)	59.8 (2,969)	
Radiation therapy- % (n)					
Yes	36.9 (1,278)	50.5 (192)	9.6 (107)	31.8 (1,577)	<0.001*
No	59.4 (2,057)	42.4 (161)	85.0 (952)	63.8 (3,170)	
Immunotherapy- % (n)					
Yes	0.7 (24)	0.0 (0)	0.6 (7)	0.6 (31)	<0.001*
No	88.3 (3,059)	84.7 (322)	82.2 (921)	86.6 (4,302)	
Surgery- % (n)					
Yes	87.0 (2,998)	52.6 (200)	81.1 (908)	82.7 (4,106)	<0.001*
No	9.8 (341)	38.7 (147)	13.3 (149)	12.8 (637)	

[a] Data not presented to avoid back calculation
[b] Cells with fewer than 6 cases have been suppressed and indicated as "<6"
[c] Percentages may not add up to 100% due to missing data

Table 2.3: CSA distribution of female cancer cases and cancer stage in Baltimore City, MD

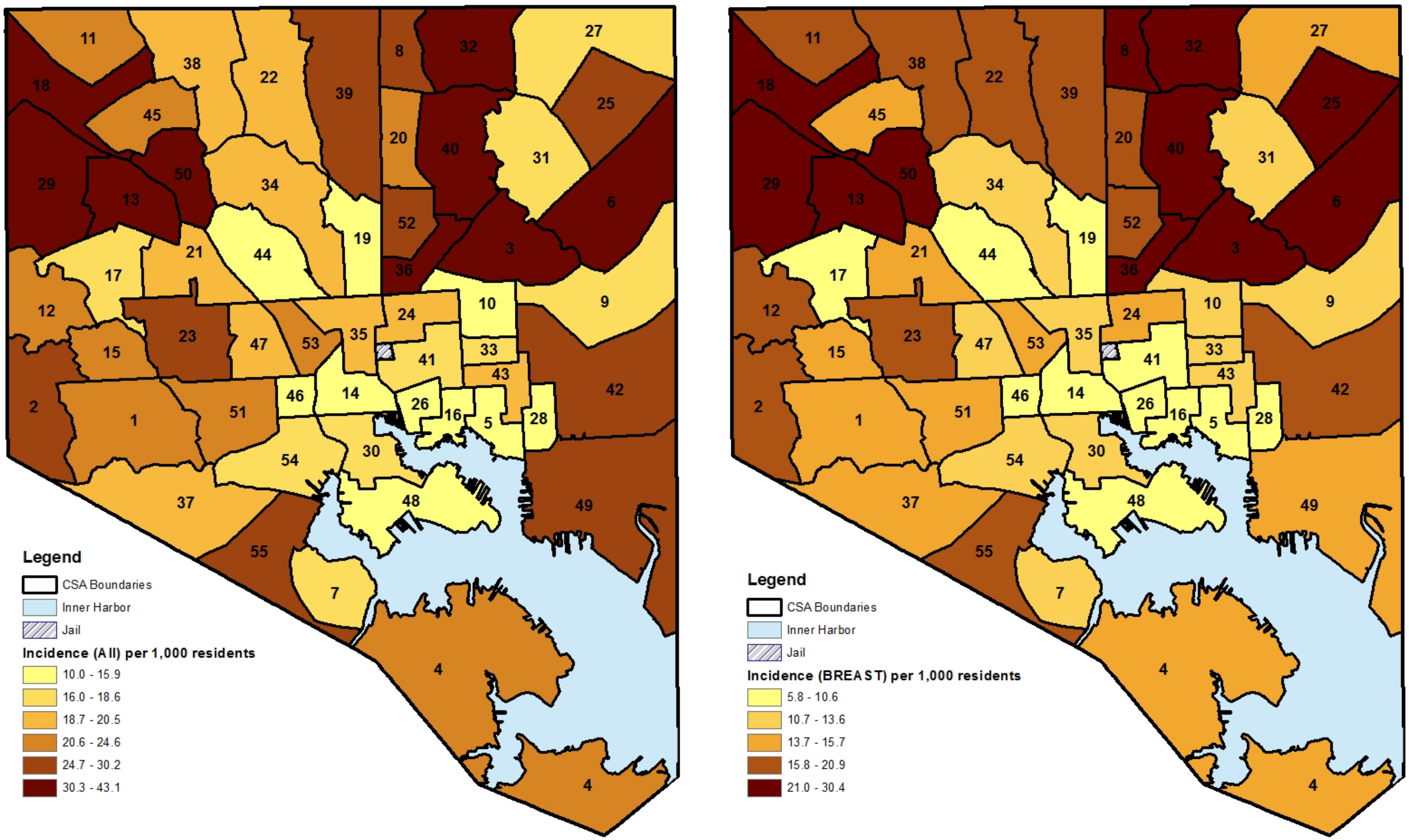
	Cases (n= 4,966)	Stage 1 (n= 2,106)	Stage 2 (n= 818)	Stage 3 (n= 1,061)	Stage 4 (n= 493)
Allendale/Irvington/S. Hilton	139	52	24	31	16
Beechfield/Ten Hills/West	88	32	14	19	11
Belair-Edison	155	70	26	35	16
Brooklyn/Curtis Bay	109	42	14	26	12
Canton	49	21	6	14	6
Cedonia/Frankford	191	81	37	41	20
Cherry Hill	52	21	12	10	7
Chinquapin Park/Belvedere	71	36	10	12	6
Claremont/Armistead	68	33	11	13	<6
Clifton-Berea	83	36	11	21	10
Cross-Country/Cheswolde	85	33	9	18	11
Dickeyville/Franklintown	26	11	<6	<6	0
Dorchester/Ashburton	126	42	15	39	14
Downtown/Seton Hill	29	12	7	<6	4
Edmonson Village	92	34	16	22	10
Fells Point	51	19	7	17	<6
Forest Park/Walbrook	93	34	11	22	12
Glen-Fallstaff	140	56	21	25	14
Greater Charles Village/Barclay	81	37	13	20	7
Greater Govans	107	51	14	21	11
Greater Mondawmin	87	30	14	22	11
Greater Roland Park/Poplar Hill	62	36	9	9	<6
Greater Rosemont	168	70	26	39	23
Greenmount East	93	34	15	23	13
Hamilton	116	51	24	25	12
Harford/Echodale	31	14	<6	8	<6
Highlandtown	124	50	31	23	9
Howard Park/West Arlington	42	18	10	7	6
Inner Harbor/Federal Hill	117	51	15	27	7
Jonestown/Oldtown	88	41	17	15	9
Lauraville	102	52	13	20	7
Loch Raven	144	65	18	25	13
Madison/East End	47	18	13	11	<6
Medfield/Hampden/Woodberry/Remington	110	54	19	14	6
Midtown	82	36	16	18	<6
Midway/Coldstream	98	45	14	20	10
Morrell Parkk/Violetville	74	33	8	16	9
Mount Washington/Coldspring	57	35	11	7	<6
North Baltimore/Guilford/Homeland	137	83	22	16	7
Northwood	144	68	22	30	12
Orangeville/E. Highlandtown	95	41	17	17	7
Patterson Park North & East	85	32	19	17	10
Penn North/Reservoir Hill	108	33	23	31	12
Perkins/Middle East	54	21	6	14	11
Pimlico/Arlington/Hilltop	101	36	15	15	16
Poppleton/The Terraces/Hollins Market	27	10	<6	<6	<6
Sandtown-Winchester/Harlem Park	126	41	26	35	16
South Baltimore	32	19	6	<6	<6
Southeastern	57	24	10	12	8
Southern Park Heights	131	60	14	32	11
Southwest Baltimore	155	61	36	31	14
The Waverlies	72	30	12	20	6
Upton/Druid Heights	78	29	14	19	7
Washington Village/Pigtown	38	14	9	7	<6
Westport/Mount Winans/Lakeland	49	18	7	14	<6

* Cells with fewer than 6 cases have been suppressed and indicated as "<6"

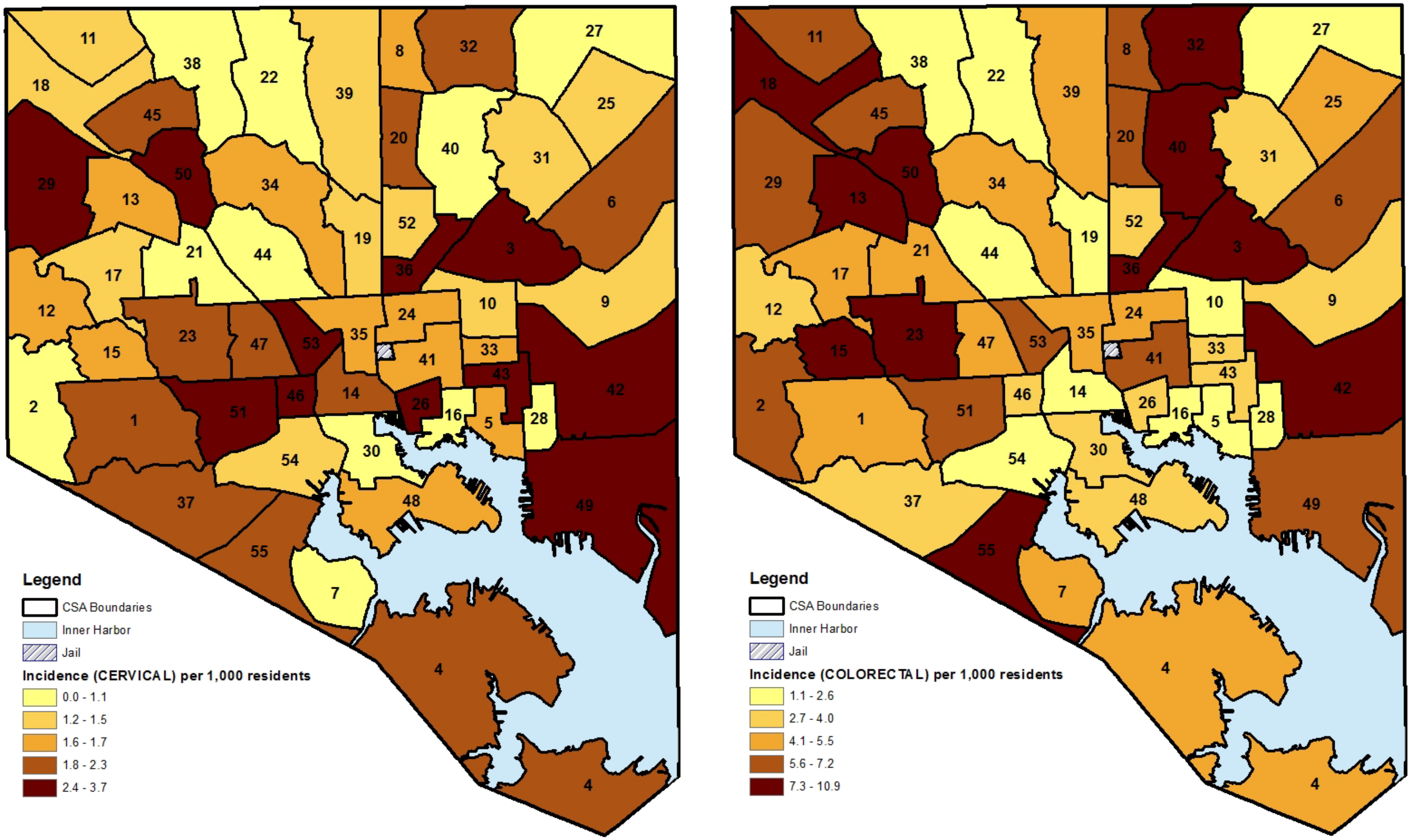
Table 2.4: Key for Community Statistical Area (CSA) map

Allendale/Irvington/S. Hilton	1	Howard Park/West Arlington	29
Beechfield/Ten Hills/West	2	Inner Harbor/Federal Hill	30
Belair-Edison	3	Lauraville	31
Brooklyn/Curtis Bay	4	Loch Raven	32
Canton	5	Madison/East End	33
Cedonia/Frankford	6	Medfield/Hampden/Woodberry/Remington	34
Cherry Hill	7	Midtown	35
Chinquapin Park/Belvedere	8	Midway/Coldstream	36
Claremont/Armistead	9	Morrell Parkk/Violetville	37
Clifton-Berea	10	Mount Washington/Coldspring	38
Cross-Country/Cheswolde	11	North Baltimore/Guilford/Homeland	39
Dickeyville/Franklintown	12	Northwood	40
Dorchester/Ashburton	13	Oldtown/Middle East	41
Downtown/Seton Hill	14	Orangeville/E. Highlandtown	42
Edmonson Village	15	Patterson Park North & East	43
Fells Point	16	Penn North/Reservoir Hill	44
Forest Park/Walbrook	17	Pimlico/Arlington/Hilltop	45
Glen-Fallstaff	18	Poppleton/The Terraces/Hollins Market	46
Greater Charles Village/Barclay	19	Sandtown-Winchester/Harlem Park	47
Greater Govans	20	South Baltimore	48
Greater Mondawmin	21	Southeastern	49
Greater Roland Park/Poplar Hill	22	Southern Park Heights	50
Greater Rosemont	23	Southwest Baltimore	51
Greenmount East	24	The Waverlies	52
Hamilton	25	Upton/Druid Heights	53
Harbor East/Little Italy	26	Washington Village/Pigtown	54
Harford/Echodale	27	Westport/Mount Winans/Lakeland	55
Highlandtown	28		

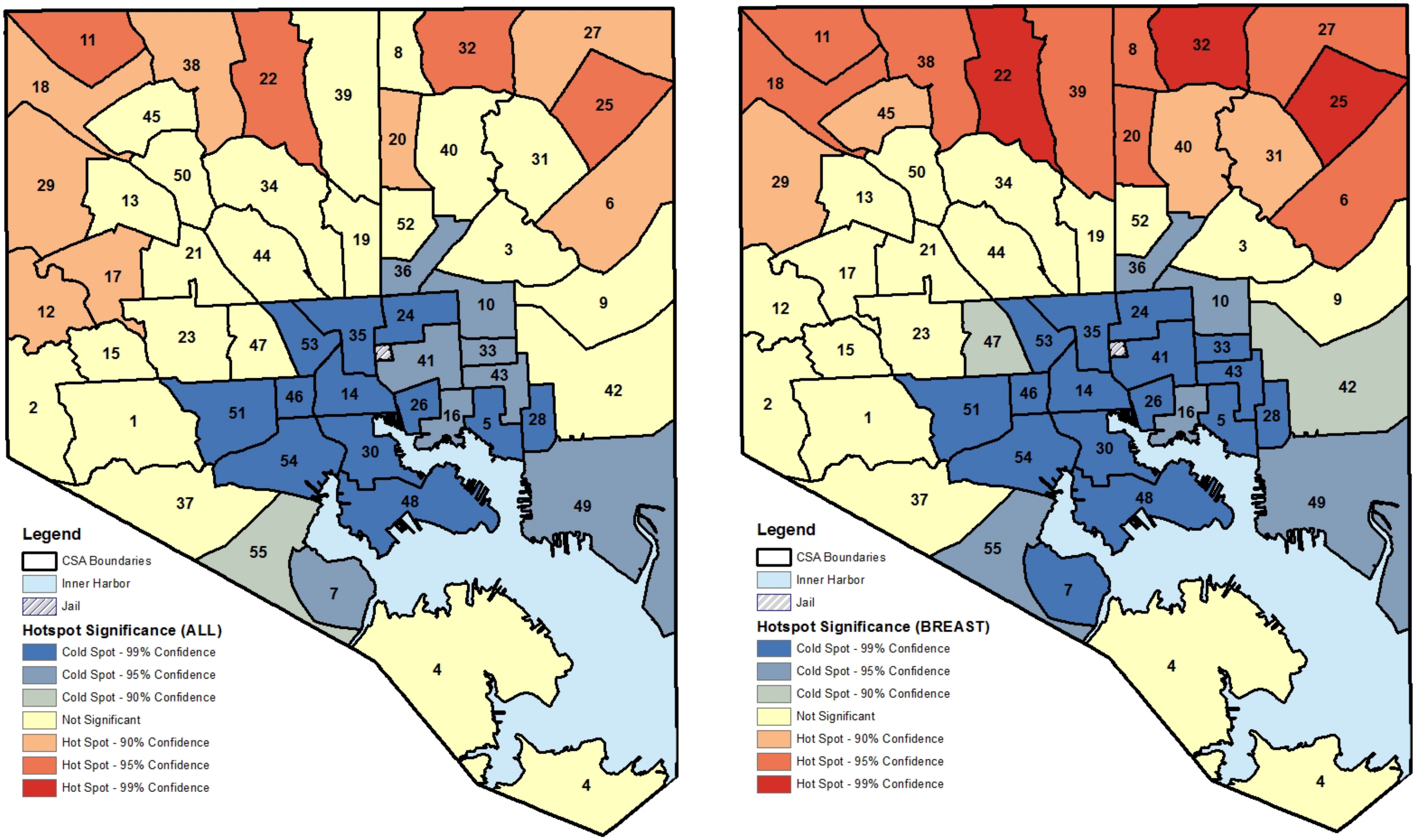
Figures 2.1a-b: CSA distribution shaded by quintile of female cancer incidence (all and breast) in Baltimore City, MD per 1,000 female residents aged 21 to 74 years



Figures 2.1c-d: CSA distribution shaded by quintile of female cancer incidence (cervical and colorectal) in Baltimore City, MD per 1,000 female residents aged 21 to 74 years



Figures 2.2a-b: Hot spot analysis and statistical significance of cancer incidence (all and breast) per 1,000 female residents aged 21 to 74 years by CSA



Figures 2.2c-d: Hot spot analysis and statistical significance of cancer incidence (cervical and colorectal) per 1,000 female residents aged 21 to 74 years by CSA

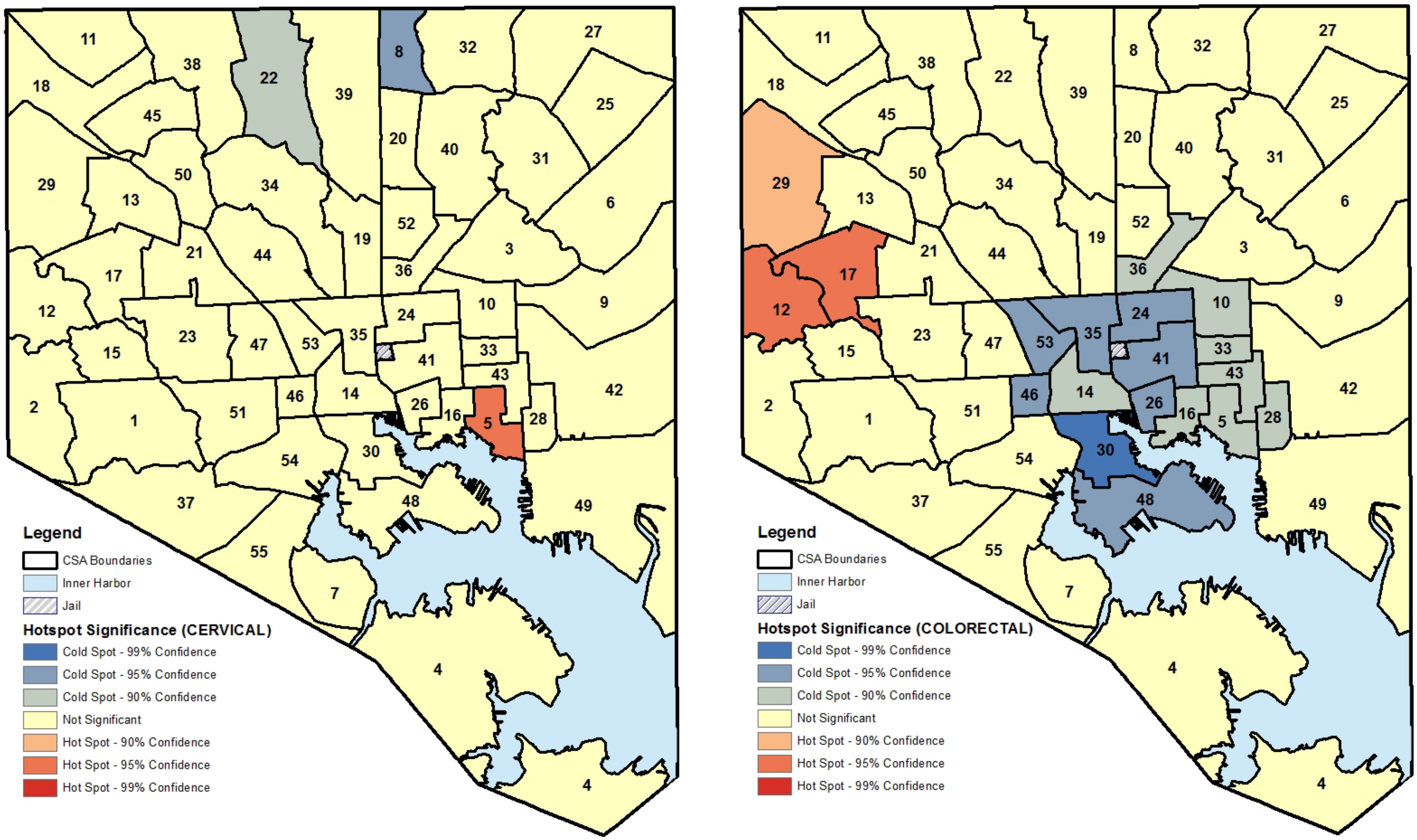
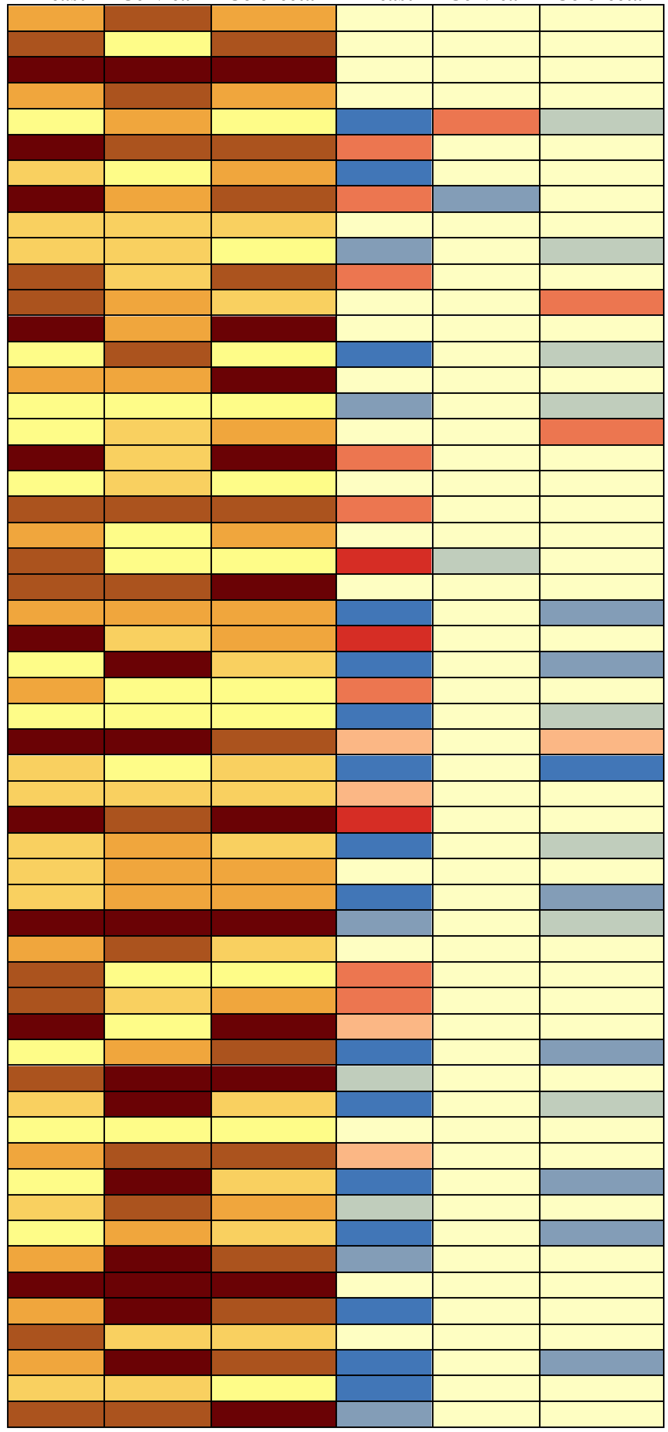


Table 2.5: Distribution of cancer site quintiles and hot/cold spatial clusters by CSA for cancer incidence

ID	CSA
1	Allendale/Irvington/S. Hilton
2	Beechfield/Ten Hills/West
3	Belair-Edison
4	Brooklyn/Curtis Bay
5	Canton
6	Cedonia/Frankford
7	Cherry Hill
8	Chinquapin Park/Belvedere
9	Claremont/Armistead
10	Clifton-Berea
11	Cross-Country/Cheswolde
12	Dickeyville/Franklintown
13	Dorchester/Ashburton
14	Downtown/Seton Hill
15	Edmonson Village
16	Fells Point
17	Forest Park/Walbrook
18	Glen-Fallstaff
19	Greater Charles Village/Barclay
20	Greater Govans
21	Greater Mondawmin
22	Greater Roland Park/Poplar Hill
23	Greater Rosemont
24	Greenmount East
25	Hamilton
26	Harbor East/Little Italy
27	Harford/Echodale
28	Highlandtown
29	Howard Park/West Arlington
30	Inner Harbor/Federal Hill
31	Lauraville
32	Loch Raven
33	Madison/East End
34	Medfield/Hampden/Woodberry/Remington
35	Midtown
36	Midway/Coldstream
37	Morrell Parkk/Violetville
38	Mount Washington/Coldspring
39	North Baltimore/Guilford/Homeland
40	Northwood
41	Oldtown/Middle East
42	Orangeville/E. Highlandtown
43	Patterson Park North & East
44	Penn North/Reservoir Hill
45	Pimlico/Arlington/Hilltop
46	Poppleton/The Terraces/Hollins Market
47	Sandtown-Winchester/Harlem Park
48	South Baltimore
49	Southeastern
50	Southern Park Heights
51	Southwest Baltimore
52	The Waverlies
53	Upton/Druid Heights
54	Washington Village/Pigtown
55	Westport/Mount Winans/Lakeland

Quintile Distribution: Breast, Cervical, Colorectal — **Cluster Spot**: Breast, Cervical, Colorectal



Lowest quintile — Highest quintile

Quintile Key

Cold Spot- 99% CI — Not Significant — Hot Spot- 99% CI

Cluster Key

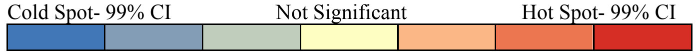
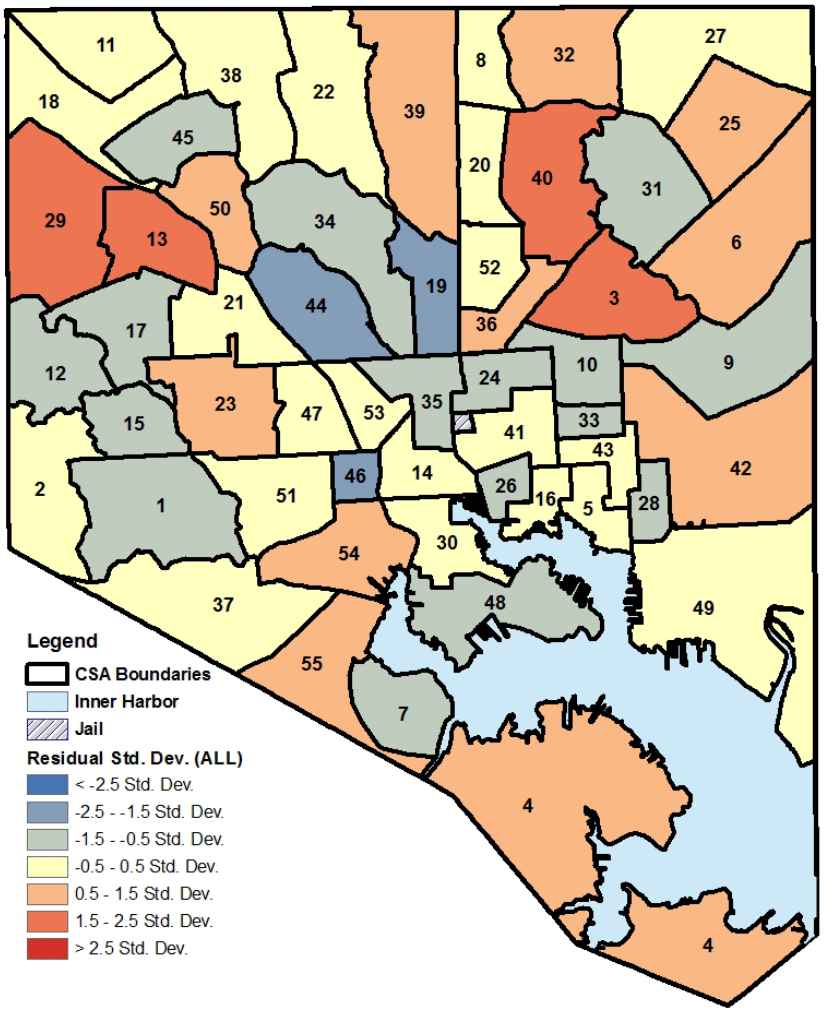


Table 2.6: Ordinary Least Squares regression models for cancer incidence by cancer site and candidate neighborhood-level covariates

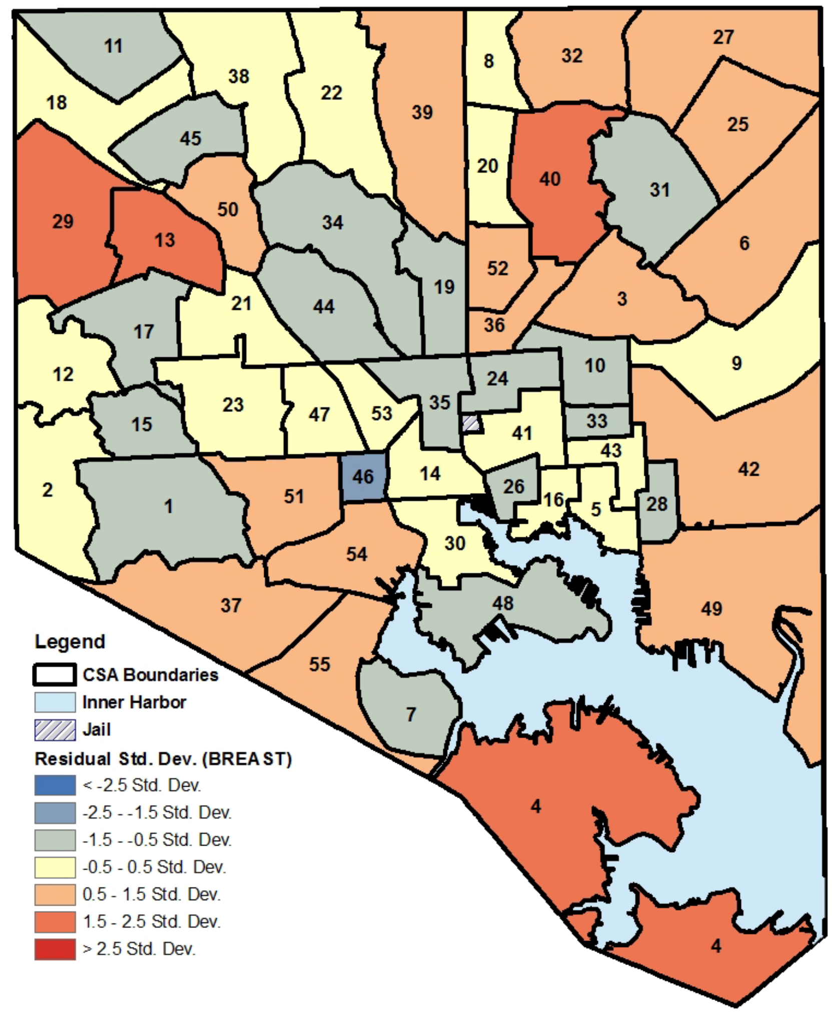
	All Cancers			Breast			Cervical			Colorectal	
	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value
Unadjusted											
Females 50-74 yrs	-0.0001	0.945		0.0001	0.918		-0.0001	0.463		0.0003	0.604
% AA	0.100	0.004*		0.059	0.017*		0.006	0.126		0.031	0.001*
Racial Diversity	-0.078	0.566		-0.052	0.201		-0.002	0.787		-0.031	0.049*
Household income <25K	-0.014	0.552		-0.059	0.377		0.024	0.009*		0.020	0.441
Female headed	0.036	0.552		-0.009	0.838		0.017	0.006*		0.028	0.107
Vacants	-0.140	0.229		-0.159	0.060		0.018	0.142		0.001	0.976
Housing violations	-0.326	0.577		-0.531	0.214		0.121	0.046*		0.084	0.622
Crime	-0.043	0.009*		-0.033	0.006*		0.001	0.642		-0.011	0.023*
Domestic violence	0.014	0.853		-0.032	0.561		0.026	<0.004*		0.020	0.363
Teen births	0.013	0.716		-0.018	0.482		0.013	<0.001*		0.018	0.078
Employed	-0.024	0.816		0.050	0.506		-0.032	0.002*		-0.041	0.157
Businesses	-0.005	0.032*		-0.003	0.033*		0.000	0.999		-0.001	0.050
Voted	0.224	0.075		0.258	0.004*		0.224	0.075		0.004	0.915
Dirty streets	-0.016	0.417		-0.017	0.231		0.004	0.085		0.002	0.692
Tree coverage	0.110	0.100		0.121	0.013*		-0.017	0.013*		0.007	0.718
Neighborhood associations	0.065	0.760		-0.033	0.835		0.024	0.291		0.074	0.227
Adjusted											
Females 50-74 yrs	-0.005	0.012*	Females 50-74 yrs	-0.003	0.058	Females 50-74 years	0.0001	0.748	Females 50-74 yrs	-0.0001	0.894
% AA	0.144	<0.001*	% AA	0.102	<0.001*	Female headed	-0.011	0.470	% AA	0.022	0.025*
Crime	-0.169	<0.001*	Crime	-0.100	0.005*	Housing violations	-0.036	0.728	Crime	-0.008	0.084
Businesses	0.017	0.002*	Businesses	0.011	0.011*	Domestic violence	0.023	0.278			
			Voted	0.222	0.029*	Teen births	-0.0001	0.982			
			Tree coverage	-0.014	0.792	Employed	-0.012	0.600			
						Tree coverage	0.006	0.502			
						Neighborhood associations	0.032	0.279			
R-squared	0.341			0.349			0.108			0.108	

* Statistically significant

Figures 2.3a-b: Spatial output of final OLS models for cancer incidence (all and breast)

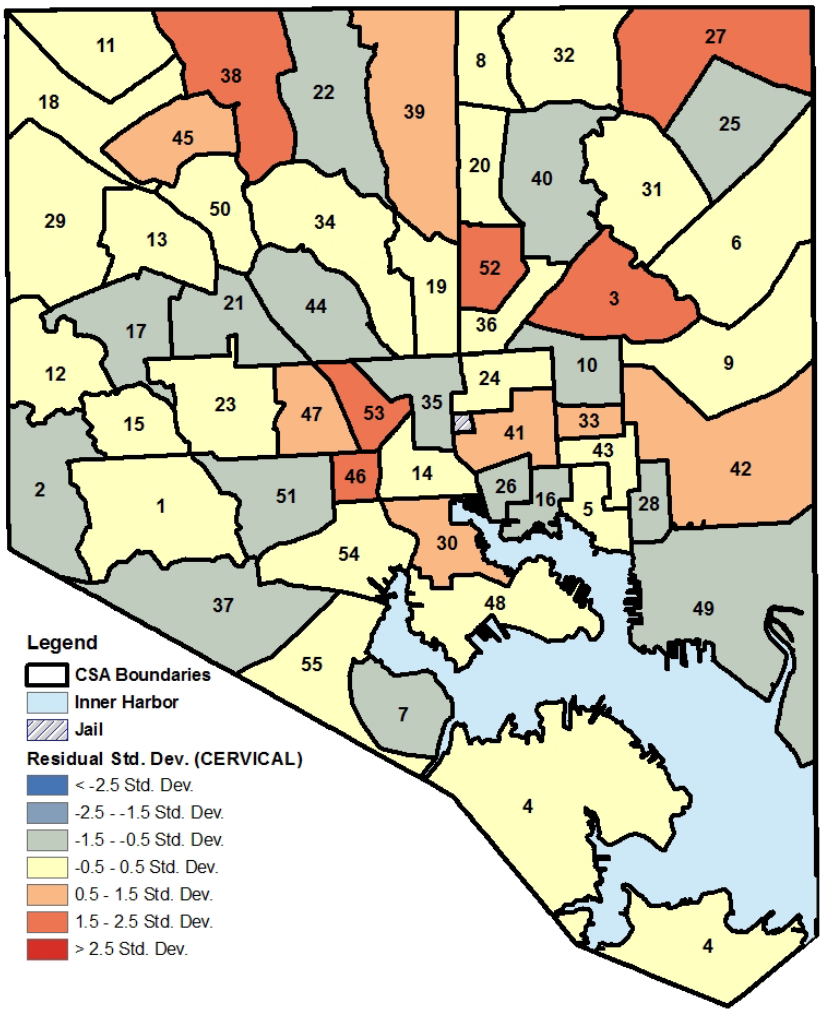


[a] Overall cancer model: Females (50-74 years), % African-American, crime, total businesses

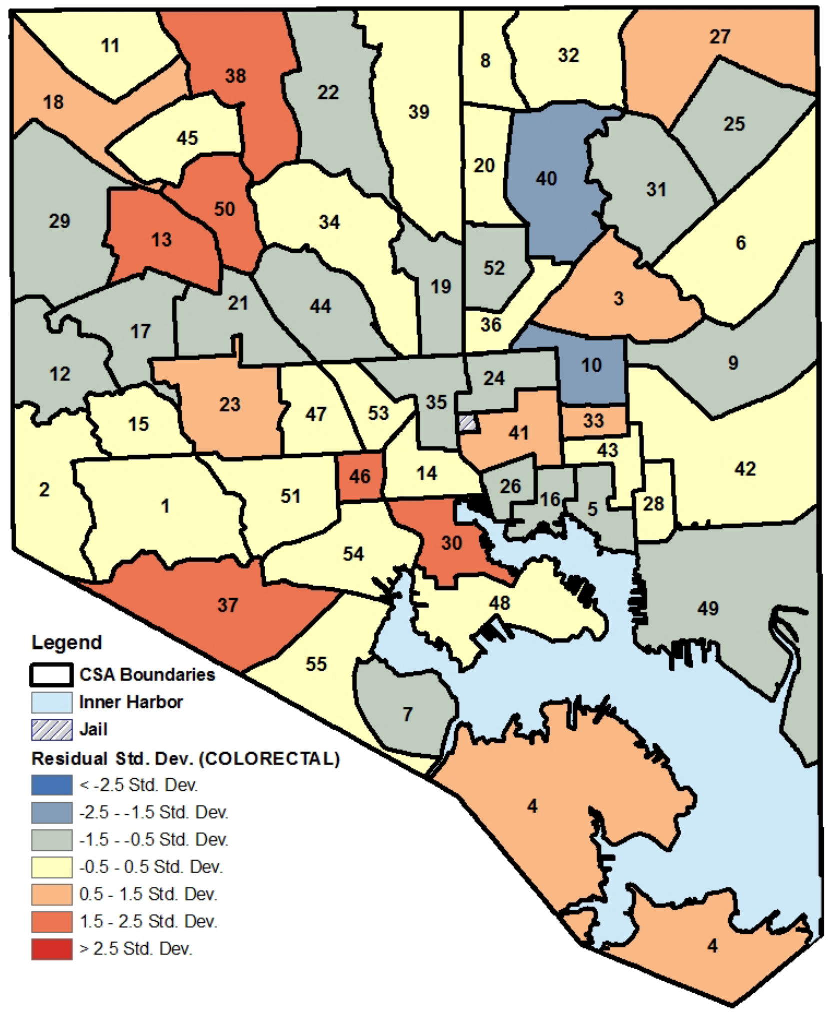


[b] Breast cancer model: Females (50-74 years), % African-American, crime, total businesses, tree coverage

Figures 2.3c-d: Spatial output of final models for ordinary least squares regression for cancer incidence (cervical and colorectal)

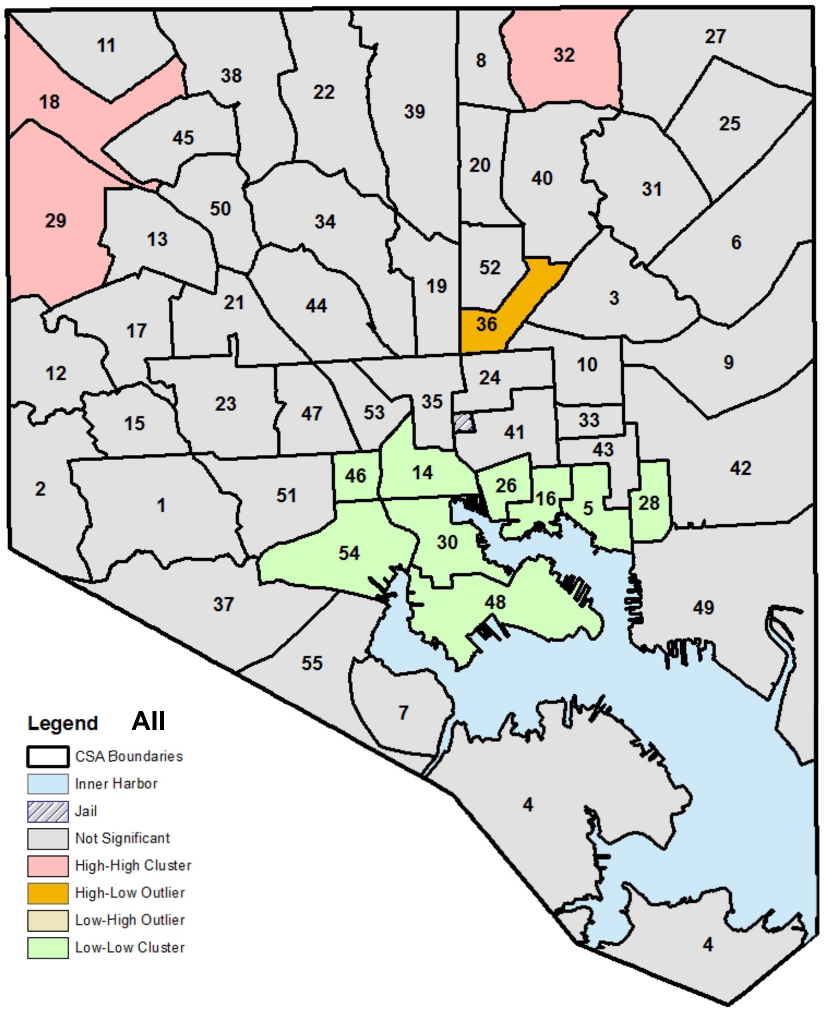


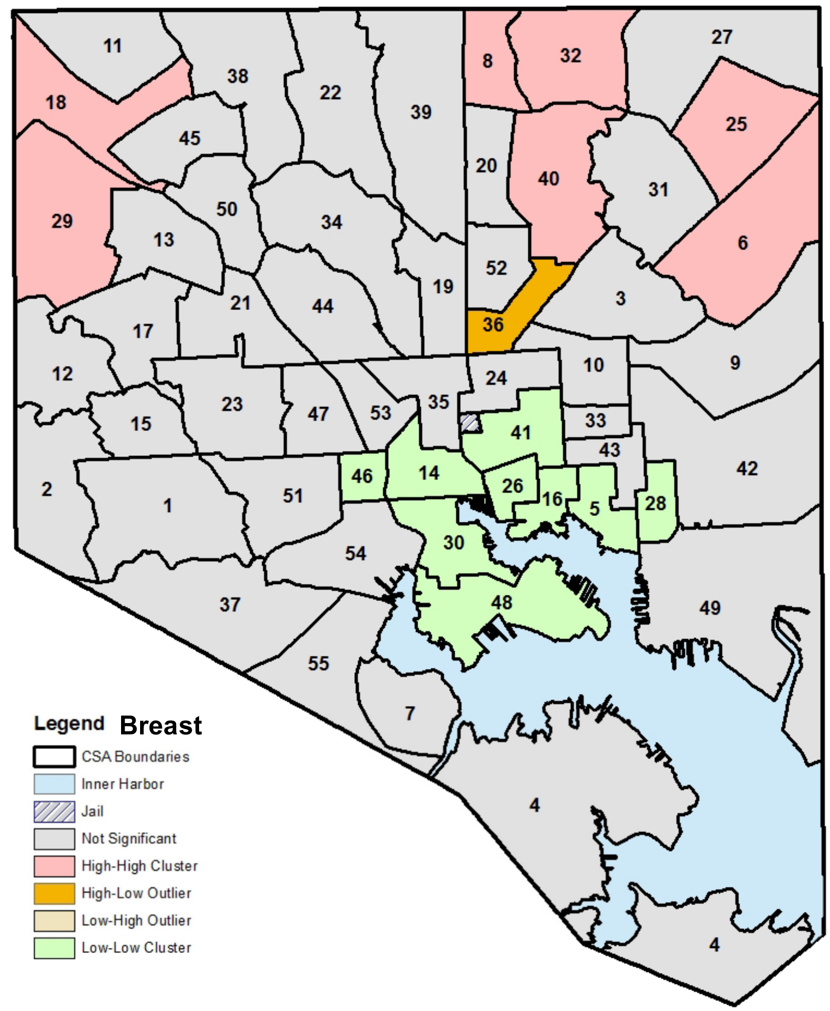
[c] Cervical cancer model: Female-headed households, housing violations, domestic violence, teen births, % employed, tree coverage, neighborhood associations



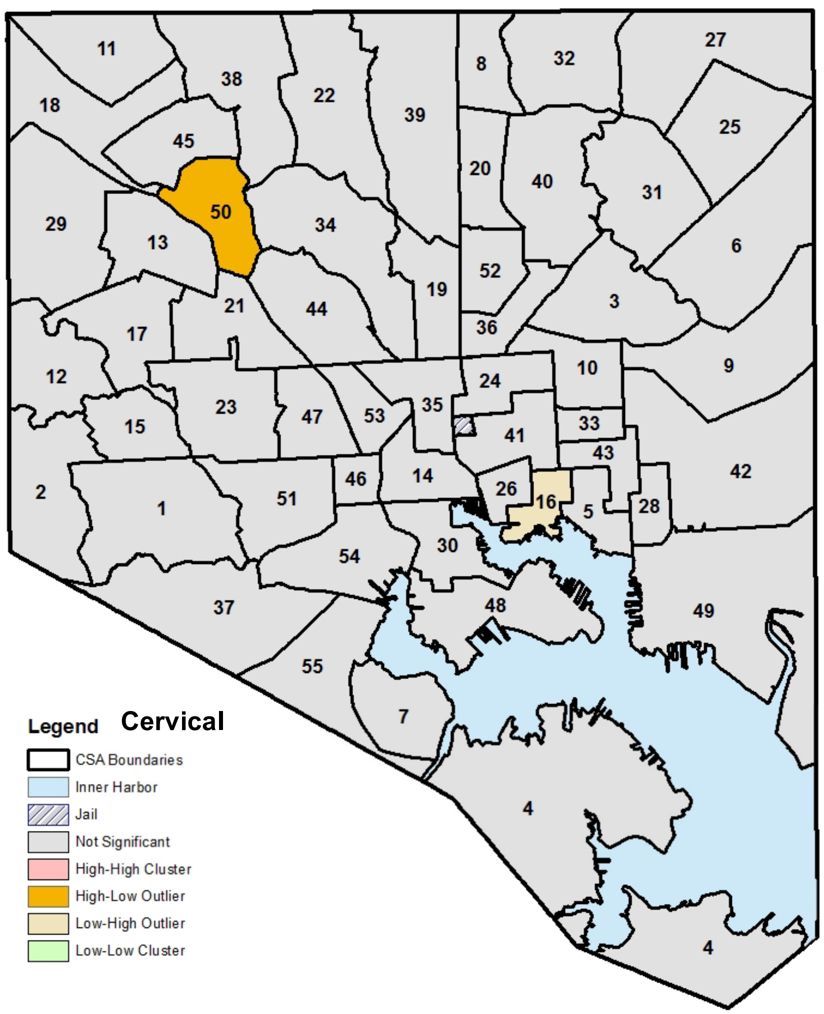
[d] Colorectal cancer model: Females (50-74 years), % African-American, crime

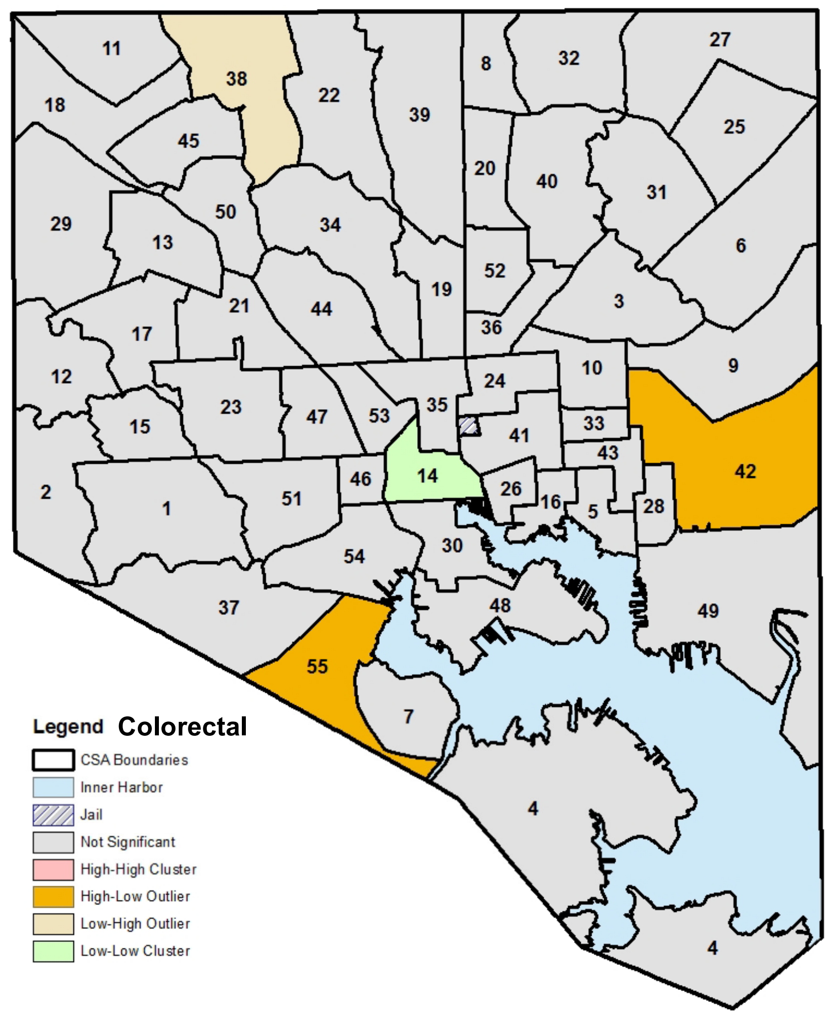
Figures 2.4a-b: Local Moran's I analysis for cancer incidence (all and breast)



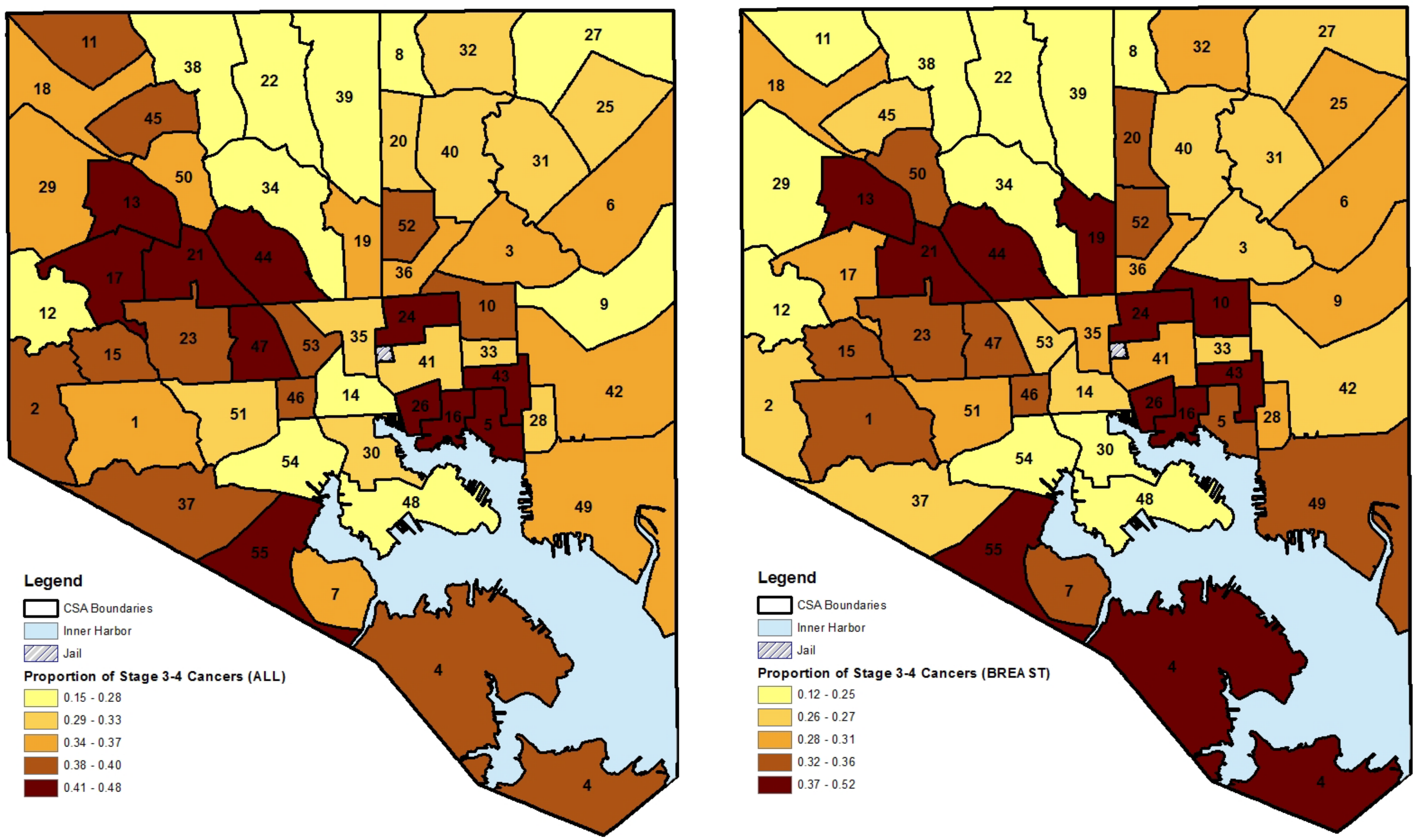


Figures 2.4c-d: Local Moran's I analysis for cancer incidence (cervical and colorectal)

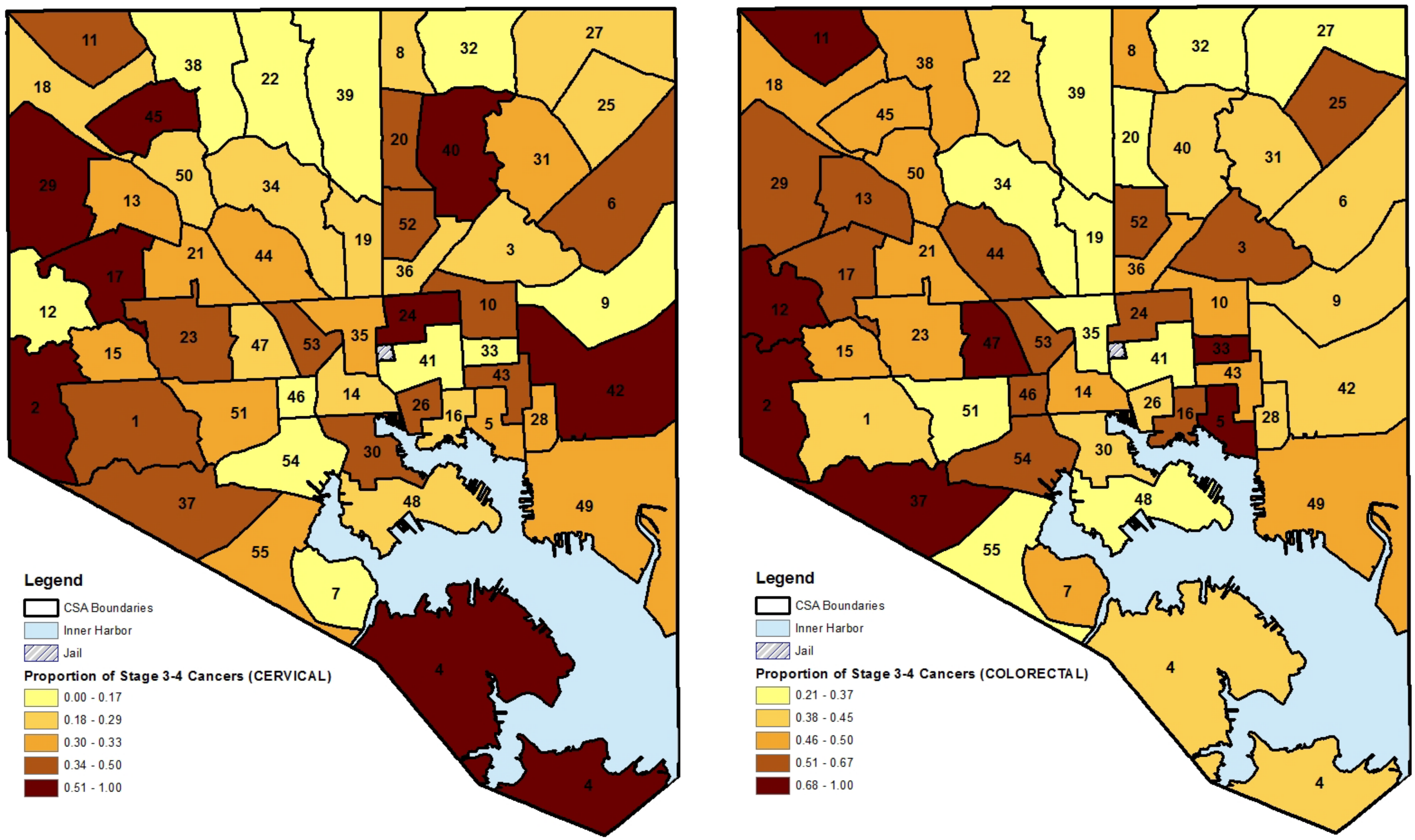




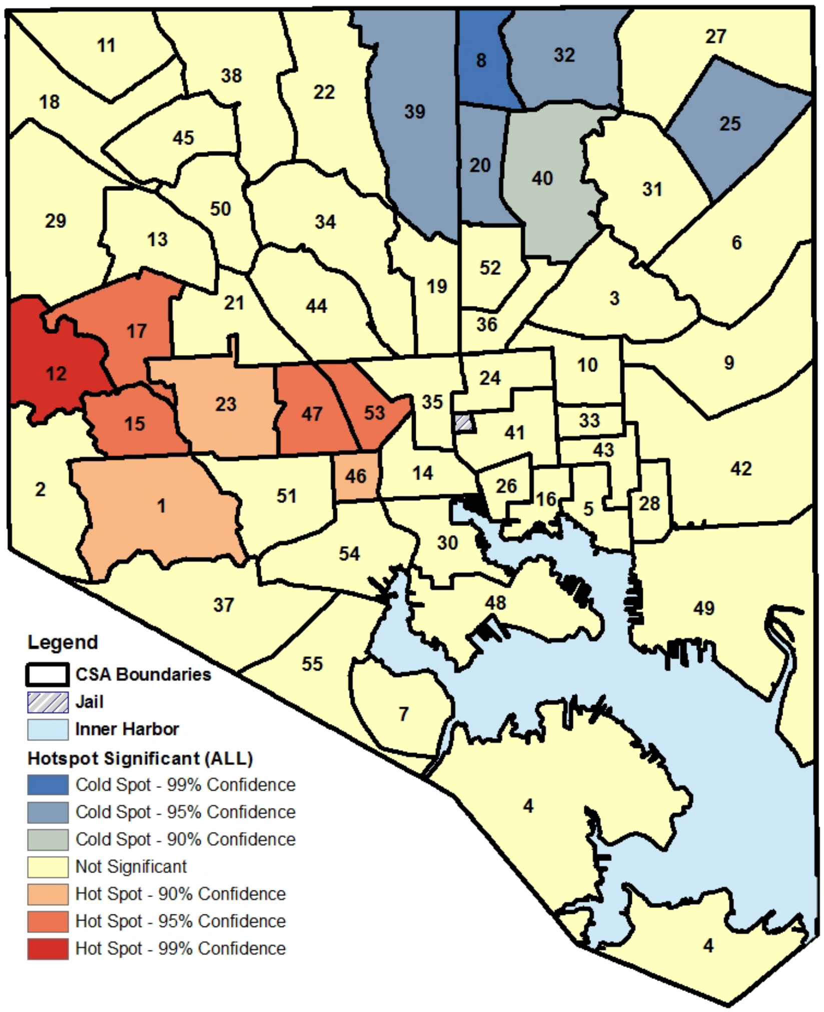
Figures 2.5a-b: CSA distribution shaded by quintile of stage 3 and 4 cancers over total cancers diagnosed (all and breast) in Baltimore City, MD

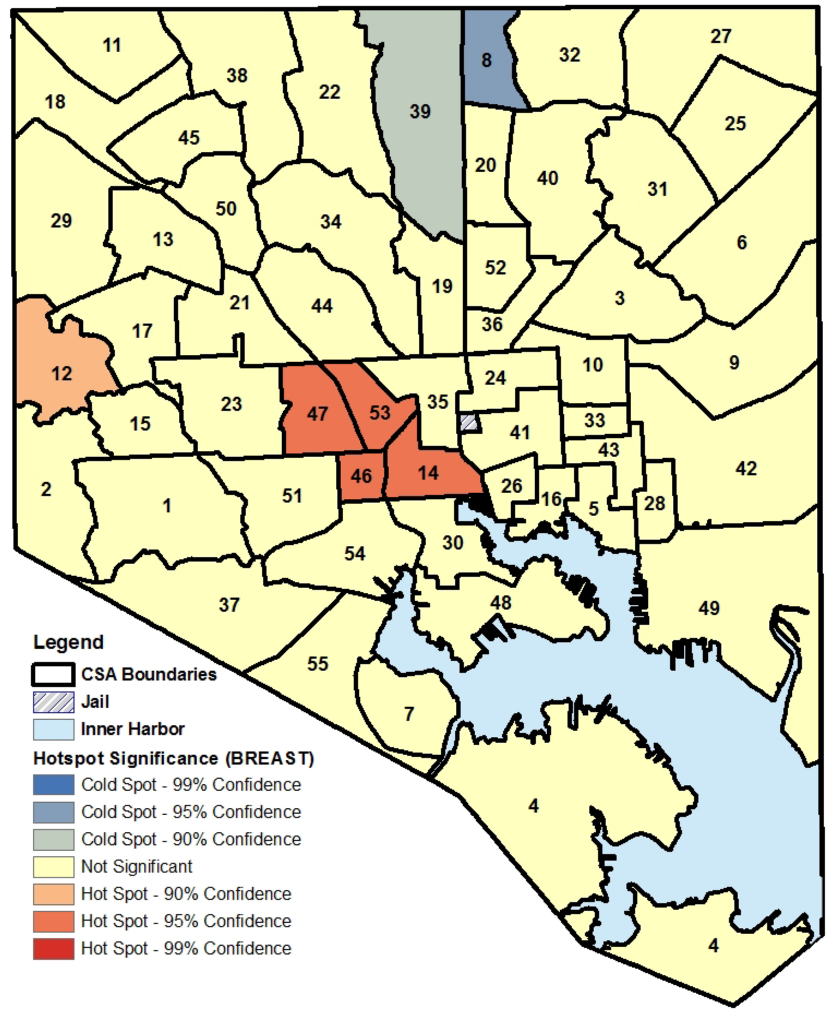


Figures 2.5c-d: CSA distribution shaded by quintile of stage 3 and 4 cancers over total cancers diagnosed (cervical and colorectal) in Baltimore City, MD



Figures 2.6a-b: Hot spot analysis and statistical significance of cancer stage (all and breast) per 1,000 female residents aged 21 to 74 years by CSA





Figures 2.6c-d: Hot spot analysis and statistical significance of cancer stage (cervical and colorectal) per 1,000 female residents aged 21 to 74 years by CSA

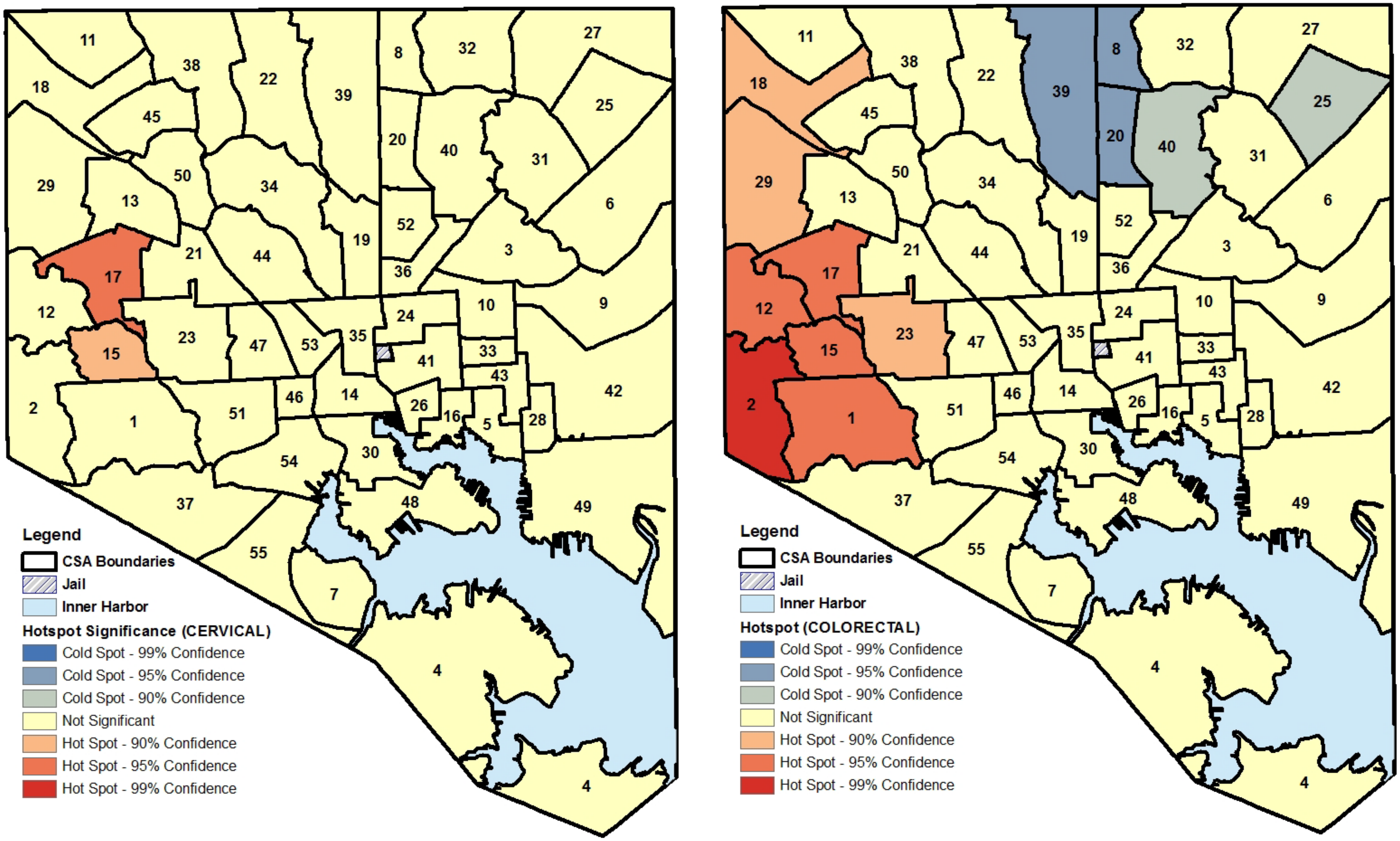
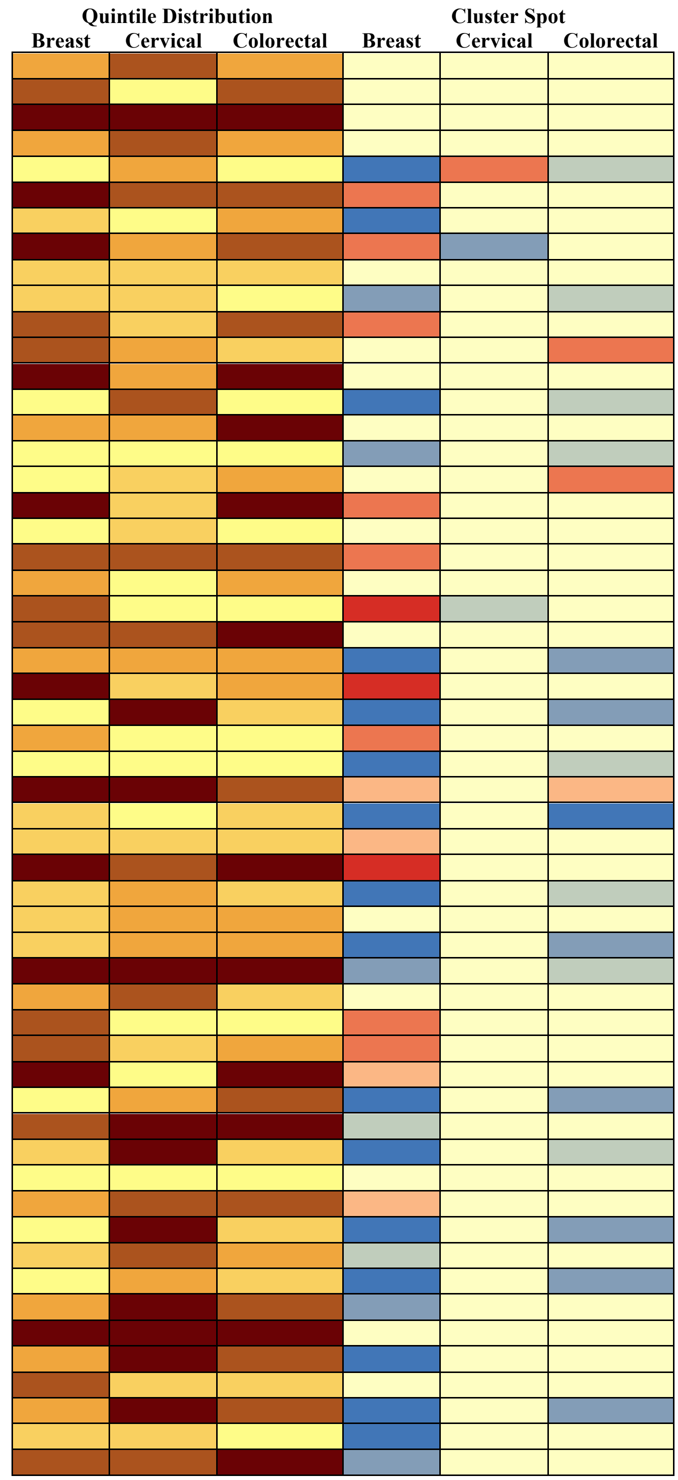


Table 2.7: Distribution of cancer site quintiles and hot/cold spatial clusters by CSA for cancer stage

ID	CSA
1	Allendale/Irvington/S. Hilton
2	Beechfield/Ten Hills/West
3	Belair-Edison
4	Brooklyn/Curtis Bay
5	Canton
6	Cedonia/Frankford
7	Cherry Hill
8	Chinquapin Park/Belvedere
9	Claremont/Armistead
10	Clifton-Berea
11	Cross-Country/Cheswolde
12	Dickeyville/Franklintown
13	Dorchester/Ashburton
14	Downtown/Seton Hill
15	Edmonson Village
16	Fells Point
17	Forest Park/Walbrook
18	Glen-Fallstaff
19	Greater Charles Village/Barclay
20	Greater Govans
21	Greater Mondawmin
22	Greater Roland Park/Poplar Hill
23	Greater Rosemont
24	Greenmount East
25	Hamilton
26	Harbor East/Little Italy
27	Harford/Echodale
28	Highlandtown
29	Howard Park/West Arlington
30	Inner Harbor/Federal Hill
31	Lauraville
32	Loch Raven
33	Madison/East End
34	Medfield/Hampden/Woodberry/Remington
35	Midtown
36	Midway/Coldstream
37	Morrell Parkk/Violetville
38	Mount Washington/Coldspring
39	North Baltimore/Guilford/Homeland
40	Northwood
41	Oldtown/Middle East
42	Orangeville/E. Highlandtown
43	Patterson Park North & East
44	Penn North/Reservoir Hill
45	Pimlico/Arlington/Hilltop
46	Poppleton/The Terraces/Hollins Market
47	Sandtown-Winchester/Harlem Park
48	South Baltimore
49	Southeastern
50	Southern Park Heights
51	Southwest Baltimore
52	The Waverlies
53	Upton/Druid Heights
54	Washington Village/Pigtown
55	Westport/Mount Winans/Lakeland



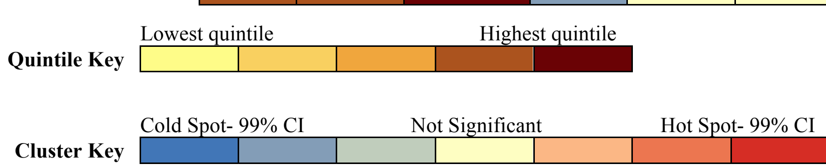
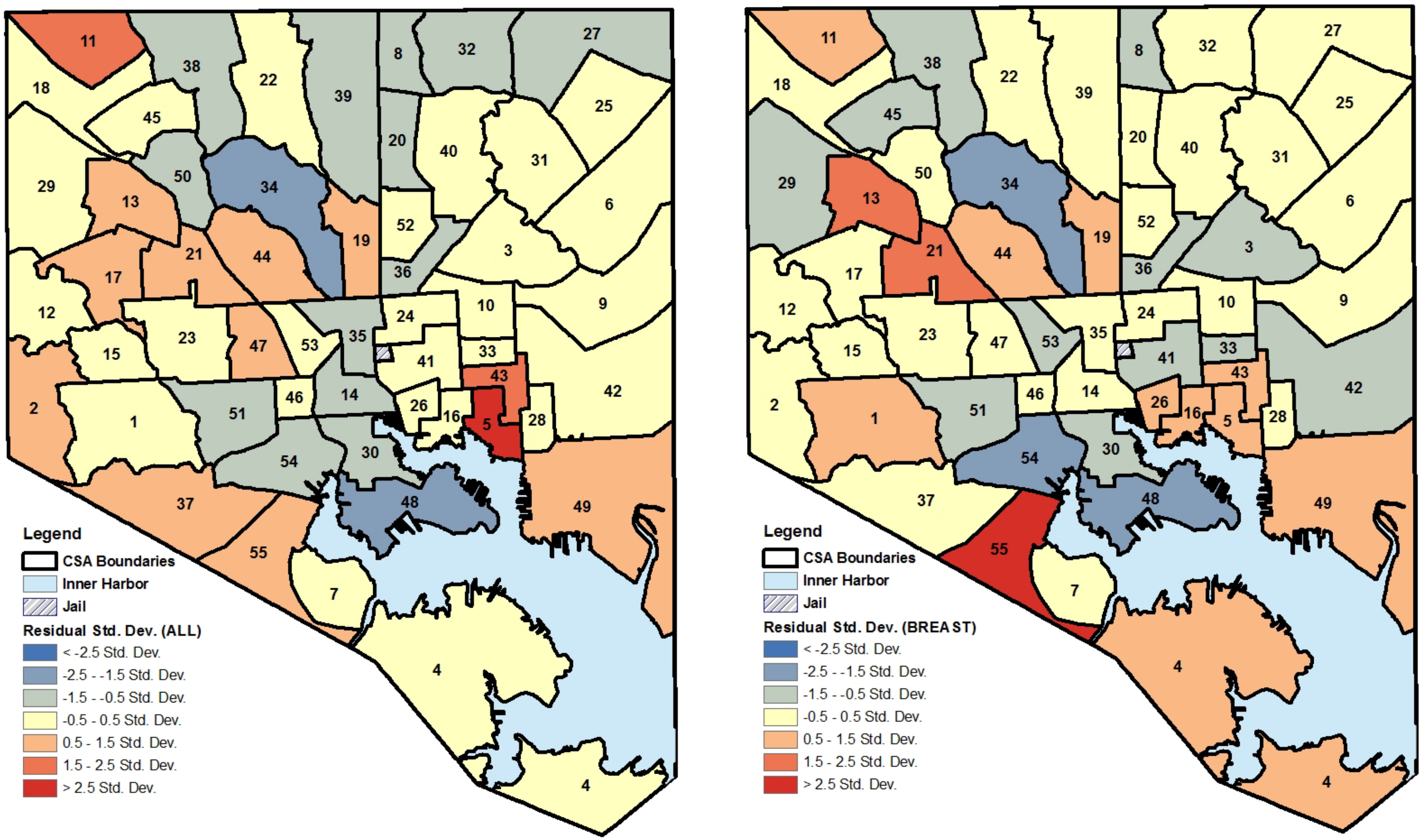


Table 2.8: Ordinary Least Squares regression models for cancer stage by cancer site and candidate neighborhood-level covariates

	All Cancers			Breast		Cervical			Colorectal	
	Coefficient	p-value		Coefficient	p-value	Coefficient	p-value		Coefficient	p-value
Unadjusted										
Females										
50-74 yrs	0.00001	0.625		0.00001	0.643	0.0001	0.096		-0.0001	0.045*
% AA	0.001	<0.001*		0.001	0.002*	0.001	0.235		0.001	0.160
Racial Diversity	-0.001	0.090		0.000	0.737	-0.001	0.483		-0.002	0.045*
Household income										
<25K	0.002	0.003*		0.003	0.005*	-0.001	0.803		0.002	0.207
Female headed	0.002	<0.001		0.002	<0.001	0.000	0.835		0.001	0.256
Vacants	0.002	0.038*		0.002	0.075	-0.002	0.540		0.002	0.444
Housing violations	0.019	<0.001*		0.019	0.001*	0.015	0.360		0.011	0.386
Crime	0.000	0.861		0.000	0.730	0.000	0.647		0.000	0.828
Domestic violence	0.002	0.001*		0.002	0.001*	0.000	0.854		0.002	0.142
Teen births	0.001	<0.001*		0.001	<0.001	0.001	0.450		0.001	0.259
Employed	-0.002	0.012*		-0.003	0.010*	0.000	0.891		0.000	0.890
Businesses	0.000	0.136		0.000	0.381	0.000	0.824		0.000	0.195
Voted	-0.432	0.093		-0.004	0.006*	-0.001	0.874		0.000	0.873
Dirty streets	0.000	0.064		0.000	0.151	0.000	0.595		0.000	0.349
Tree coverage	-0.001	0.038*		-0.002	0.019*	0.000	0.903		0.002	0.173
Neighborhood associations	0.004	0.054		0.004	0.093	0.006	0.273		-0.001	0.847
Adjusted										
Females			Females					Females		
50-74 yrs	0.00002	0.185	50-74 yrs	0.00001	0.293			50-74 yrs	-0.0001	0.024*
Racial Diversity	-0.0003	0.607	Female headed	0.002	0.066			Racial Diversity	-0.002	0.024*
Female headed	0.001	0.229	Housing violations	0.007	0.392					
Vacant	-0.004	0.023*	Domestic violence	-0.002	0.198					
Housing violations	0.021	0.007*	Teen births	0.001	0.019*					
Domestic violence	-0.002	0.131	Employed	0.002	0.260					
Teen births	0.002	0.001	Voted	-0.0001	0.969					
Employed	0.001	0.640	Trees	-0.001	0.360					
Tree coverage	0.001	0.257								
R-squared	0.392			0.285					0.129	

* Statistically significant

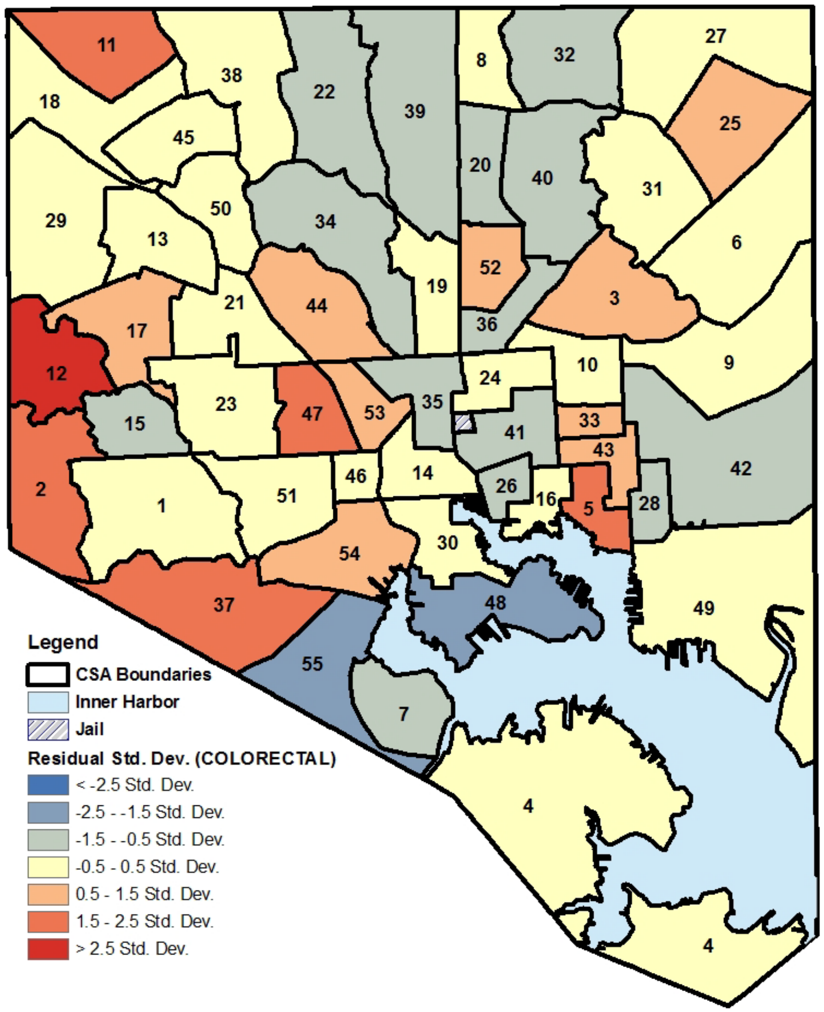
Figures 2.7a-b: Spatial output of final models for ordinary least squares regression for cancer stage (all and breast)



[a]Overall cancer model: Females (50-74 years), racial diversity, female-headed, housing violations, teen births, employed, tree coverage

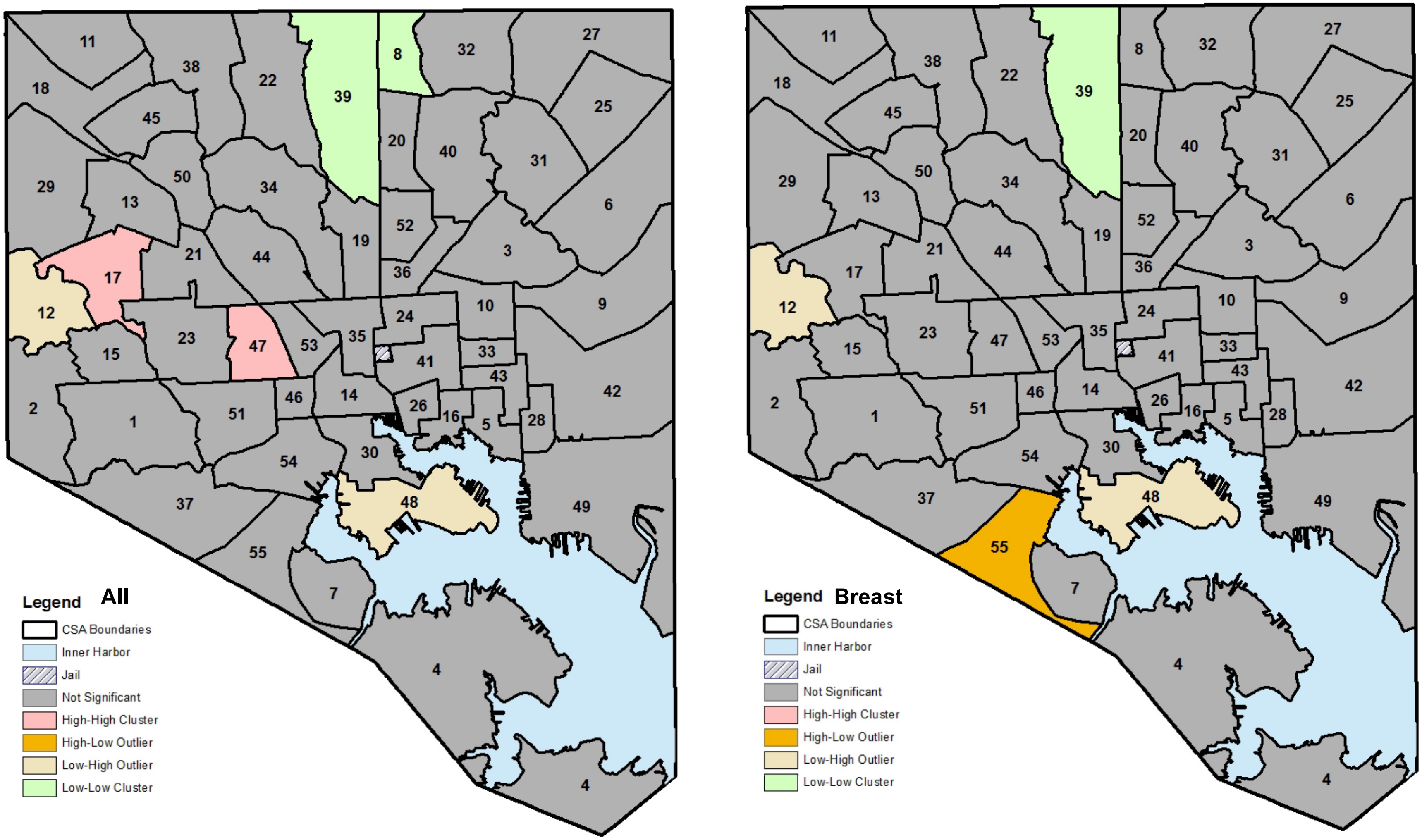
[b]Breast cancer model: Females (50-74 years), female-headed, housing violations, domestic violence, teen births, employed, voted, tree coverage

Figures 2.7c: Spatial output of final models for ordinary least squares regression for cancer stage (colorectal)

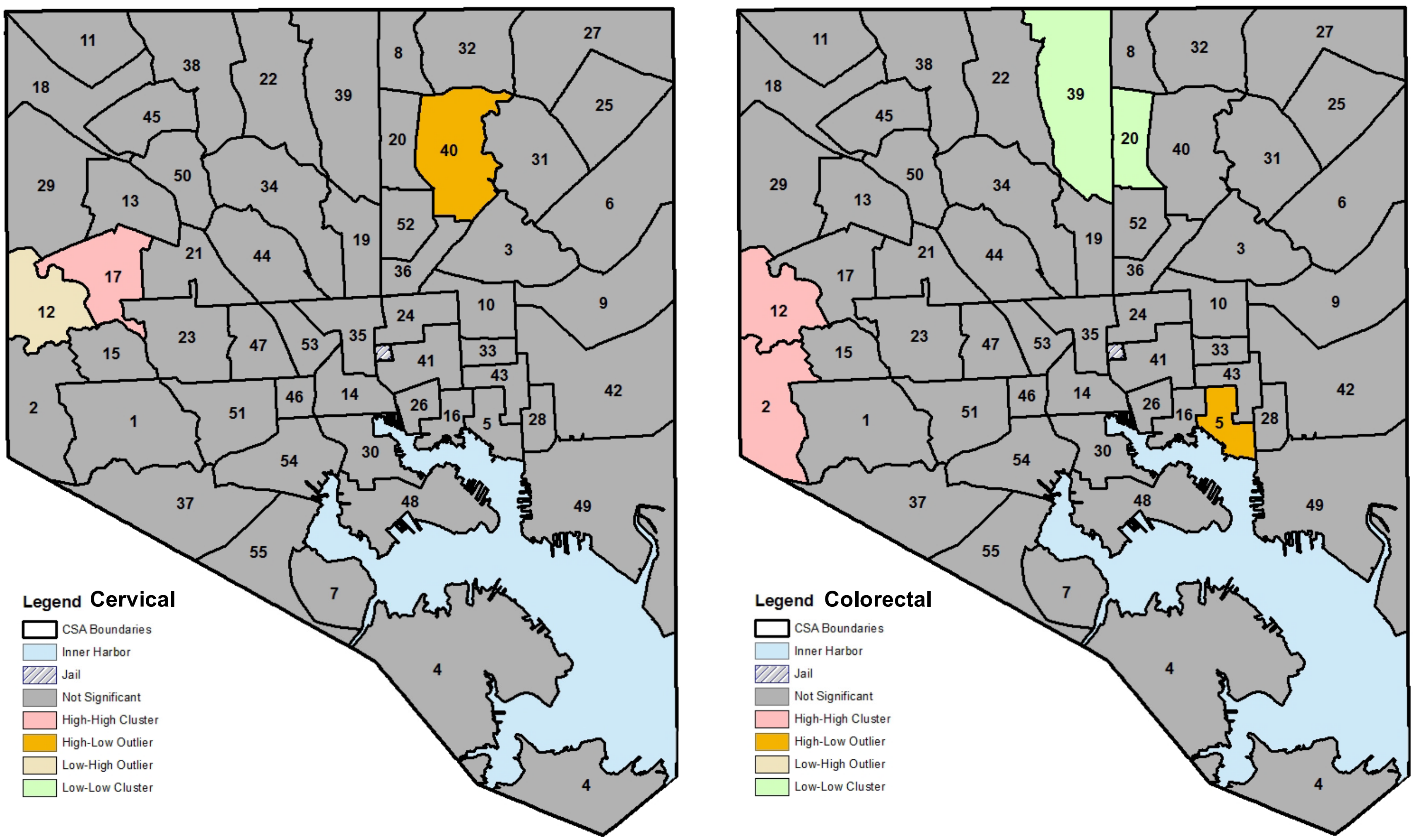


[c]Colorectal cancer model: Females (50-74 years), racial diversity

Figures 2.8a-b: Local Moran's I analysis for cancer stage (all and breast)



Figures 2.8c-d: Local Moran's I analysis for cancer stage (cervical and colorectal)



Appendix 2A Table S2.1: Definitions and years of Community Statistical Area characteristics

Domain	Characteristic	Definition	Year(s) of data
CENSUS	Females 21-49	Total number of female residents age 21 to 49 years	2000, 2010
	Females 50-74	Total number of female residents age 50 to 74 years	2000, 2010
	African-American	Percent of residents that identify themselves as being racially Black or African American (and ethnically non-Hispanic)	2000, 2010
	White	Percent of residents that identify themselves as being racially White (and ethnically non-Hispanic)	2000, 2010
	Asian	Percent of residents that identify themselves as being Asian (and ethnically non-Hispanic)	2000, 2010
	2 or more races	Percent of residents that identify themselves as being of two or more races (and non-Hispanic)	2000, 2010
	Hispanic	Percent of residents that identify their ethnicity as being Hispanic or Latino	2000, 2010
	Racial Diversity Index	Percent chance that two people picked at random within an area will be of a different race/ethnicity. The higher the value, the more racially and ethnically diverse an area	2010
	Households	Total number of households	2000, 2010
	Families	Percent of households with children under 18	2000, 2010
	Household size	Median number of persons living within a household	2000, 2010
	Female-headed	Percent of female-headed households with own children aged 18 years and younger	2000, 2010
	Median income	Median incomes earned by households	2000, 2010
	<$25K	Percent of households earning less than $25,000	2000, 2012
	$25K - $40K	Percent of households earning between $25,000 and $39,999	2000, 2012
	$40K - $60K	Percent of households earning between $40,000 and $59,999	2000, 2012
	$60K - $75K	Percent of households earning between $60,000 and $74,999	2000, 2012
	>$75K	Percent of households earning more than $75,000	2000, 2012
HOUSING & COMMUNITY DEVELOPMENT	Median home	Median price of homes sold	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Median market	Median number of days that homes listed for sale sits on the public market	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Homes sold	Percent of homes and condominiums sold that were identified as being owned by the bank (REO)	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Owner-occupied	Percent of homeowners that are the principal residents of a particular residential property	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Foreclosures	Percent of of properties where the lending company or loan servicer has filed a foreclosure proceeding	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Vacant	Percent of residential properties that have been classified as being vacant and abandoned	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Housing violations	Percent of residential properties with housing violations (excluding vacants)	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Rehab permits	Percent of residential properties that have applied for and received a permit to renovate a property where the cost of renovation will exceed $5,000	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Residential properties	Total number of residential properties	2002, 2006, 2007, 2010
	Affordabaility Index-Mortgage	Percent of households that pay more than 30% of their total household income on mortgage and other housing-related expenses out of all households	2000, 2012
	Affordabaility Index- Rent	Percent of households that pay more than 30% of their total household income on rent and related expenses out of all households	2000, 2012
	Evictions	Total number of rental evictions per 1,000 residents	2002, 2006, 2007

Appendix 2A Table S2.1: Definitions and years of Community Statistical Area characteristics (continued)

Domain	Characteristic	Definition	Year(s) of data
CRIME & SAFETY	Crime	Total number of Part 1 crime incidents per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Violent crime	Total number of violent crime incidents per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Domestic violence	Total number of calls to emergency 911 for domestic violence per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Juvenile arrests	Total number of persons aged 10-17 years arrested per 1,000 juveniles	2002, 2003, 2006, 2007, 2008, 2009
	Juvenile violence	Total number of persons aged 10-17 years arrested for violent offenses per 1,000 juveniles	2002, 2003, 2006, 2007, 2008, 2009
	Juvenile drugs	Total number of persons aged 10-17 years arrested for drug-related offenses per 1,000 juveniles	2002, 2003, 2006, 2007, 2008, 2009
	Juvenile priors	Total number of persons aged 10-17 years arrested with at least one prior offense per 1,000 juveniles	2002, 2003, 2006, 2007, 2008, 2009
EDUCATION & YOUTH	1st-5th grader ever	Total number of children who have registered for and attend 1st-5th grade at a public school at any point during the school year	2007, 2008, 2009, 2010
	6th-8th grade ever	Total number of children who have registered for and attend 6th-9th grade at a public school at any point during the school year	2007, 2008, 2009, 2010
	9th-12th grade ever	Total number of children who have registered for and attend 9th-12th grade at a public school at any point during the school year	2007, 2008, 2009, 2010
	1st-5th enrolled	Total number of children who have registered for and attend 1st-5th grade at a public school as of September 30th	2006, 2007
	6th-8th enrolled	Total number of children who have registered for and attend 6th-8th grade at a public school as of September 30th	2006, 2007
	9th-12th enrolled	Total number of children who have registered for and attend 9th-12th grade at a public school as of September 30th	2006, 2007
	1st-5th absent	Percent of 1st-5th grade students that were recognized as being absent from public school 20 or more days	2006, 2007, 2008, 2009, 2010
	6th-8th absent	Percent of 6th-8th grade students that were recognized as being absent from public school 20 or more days	2006, 2007, 2008, 2009, 2010
	9th-12th absent	Percent of 9th-12th grade students that were recognized as being absent from public school 20 or more days	2006, 2007, 2008, 2009, 2010
	Suspensions or expulsions	Percent of students of any grade level that are formally suspended or expelled for any reason during the school year	2008, 2009, 2010
	Free/reduced meals	Percent of students of any grade that are eligible for and receive free or reduced school meals	2008, 2009, 2010
	Special education	Percent of students of any grade that are eligible for and participating in public school special education programs	2008, 2009, 2010
	Dropouts	Percent of of 9th-12th graders who withdraw from public school	2006, 2007, 2008, 2009, 2010
	Completion	Percent of 12th graders in a school year that successfully completed high school	2006, 2007, 2008, 2009, 2010
	School and/or employed	Percent of residents aged 16-19 years in school and/or employed	2000, 2012

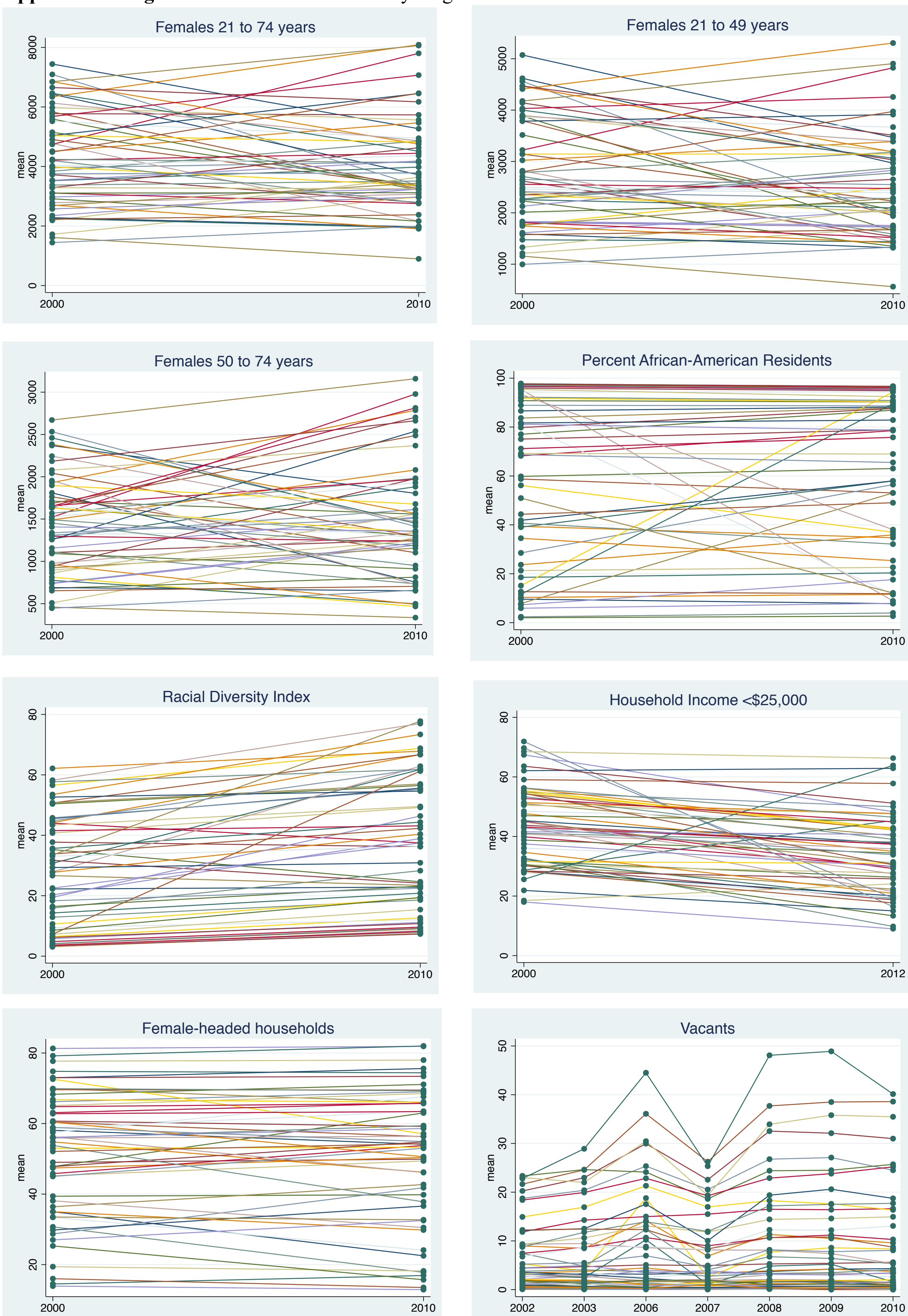
Appendix 2A Table S2.1: Definitions and years of Community Statistical Area characteristics (continued)

Domain	Characteristic	Definition	Year(s) of data
CHILDREN & FAMILY HEALTH	Teen birth	Total female teens aged 15-19 years that gave birth per 1,000 females aged 15-19 years	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Delivered at term	Percent of births delivered at term	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Birth weight	Percent of children born with a birth weight of at least 5 pounds	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Prenatal	Percent of births where the mother received prenatal care during the first trimester of the pregnancy	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Lead test	Total number of children aged 0-6 years who are tested for the presence of blood lead in a calendar year	2006, 2007, 2008, 2009, 2010
	Elevated lead	Percent of children aged 0-6 years that are found to either have elevated blood lead levels or lead poisoning	2006, 2007, 2008, 2009, 2010
WORKFORCE & ECONOMIC DEVELOPMENT	Employed	Percent of persons aged 16-64 years formally employed or self-employed	2000, 2012
	Unemployed	Percent of persons aged 16-64 not working out of all persons (not just those in the labor force)	2000, 2012
	Not in labor force	Percent of persons who are not in the labor force out of all persons aged 16-64 years	2000, 2012
	Unemployment rate	Percent of persons aged 16-64 years that are in the labor force (and are looking for work) but are not currently working	2000, 2012
	Commercial properties	Total number of commercial properties	2002, 2003, 2006, 2007, 2008, 2010
	Commercial rehab	Percent of properties that are investing above $5,000 within their current establishment	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Commercial vacants	Total number of commercial properties that are classified as vacant and abandoned at year's end	2002, 2006, 2007
	Businesses	Total number of businesses (both for-profit and non-profit)	2007, 2008, 2009, 2010
	Employees	Total number of persons employed by businesses (both for-profit and non-profit)	2007, 2008, 2009, 2010
	Small businesses	Total number of businesses (both for-profit and non-profit) that report having less than 50 persons employed	2007, 2008, 2009, 2010
	Mid-sized businesses	Total number of businesses (both for-profit and non-profit) that report having 50-100 persons employed	2007, 2008, 2009, 2010
SUSTAINABILITY	Registered voters	Percent of persons aged 18 years and over registered to vote	2002, 2004, 2006, 2010
	Registered voters- youth	Percent of persons aged 18-25 years registered to vote	2002, 2004, 2006, 2008, 2009
	Voted	Percent of persons who voted in the last general election	2002, 2004, 2006, 2010
	Voted- youth	Percent of persons aged 18-25 years who voted in the last general election	2002, 2004, 2006, 2008, 2009
	Dirty streets	Total number of service requests for dirty streets and alleys per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Clogged drains	Total number of service requests for clogged storm drains per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009, 2010
	Abandoned vehicles	Total number of abandoned vehicle reports per 1,000 residents	2002, 2006, 2007
	Rats	Total number of rat incident reports per 1,000 residents	2002, 2003, 2006, 2007, 2008, 2009
	0-14 mins travel	Percent of commuters that spend less than 15 minutes commuting to work	2000, 2012
	15-29 mins travel	Percent of commuters that spend between 15-29 minutes commuting to work	2000, 2012
	30-44 mins travel	Percent of commuters that spend between 30-44 minutes commuting to work	2000, 2012

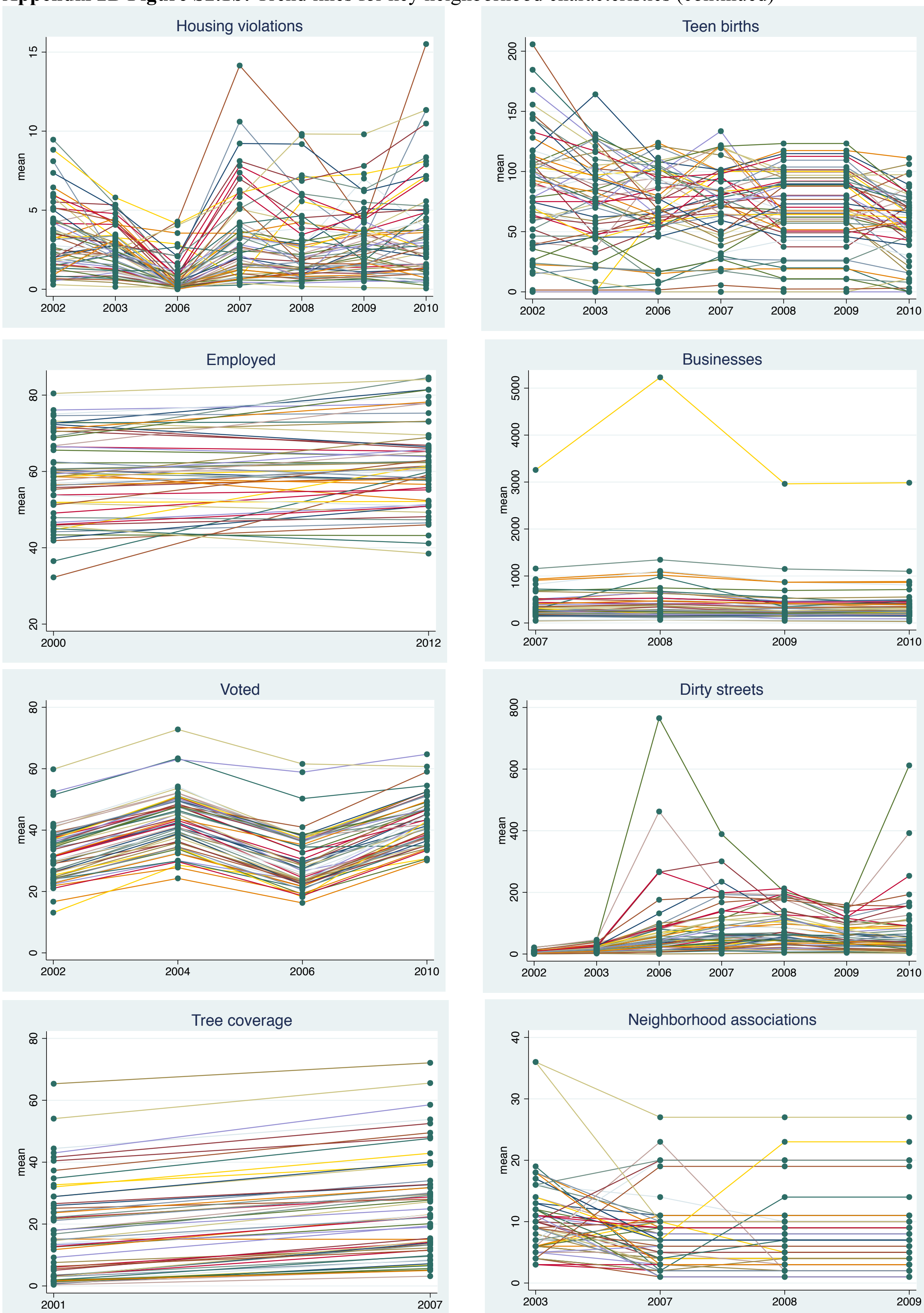
Appendix 2A Table S2.1: Definitions and years of Community Statistical Area characteristics (continued)

Domain	Characteristic	Definition	Year(s) of data
SUSTAINABILITY (continued)	>45 mins travel	Percent of commuters that spend more than 45 minutes commuting to work	2000, 2012
	Tree coverage	Percent of total land area comprised of tree canopy	2001, 2007
	Neighborhood associations	Total number of neighborhood associations and block clubs	2003, 2007, 2008, 2009
	Community development	Total number of community development corporations	2003, 2007, 2008, 2009
	Umbrella organizations	Total number of "umbrella" organizations	2003, 2007, 2008, 2009
	Parks groups	Total number of parks and environmental stewardship groups	2003, 2007, 2008, 2009
	Historic buildings	Total number of designated local historic buildings (CHAP)	2003, 2007, 2008, 2009

Appendix 2B Figure S2.1a: Trend lines for key neighborhood characteristics



Appendix 2B Figure S2.1b: Trend lines for key neighborhood characteristics (continued)



Appendix 2C Table S2.2: Characteristics of Community Statistical Areas in Baltimore City, MD

	Females 21-49 years (n)	**Females 50-74 years (n)**	**African-American (%)**	**Racial Diversity (%)**	**Female headed (%)**	**<$25K (%)**
Allendale/Irvington/S. Hilton	4,269.5	2,086.5	87.4	22.5	56.1	41.5
Beechfield/Ten Hills/West	2,077.5	1,179.5	77.0	37.0	49.5	24.5
Belair-Edison	2,785.5	1,469.0	82.0	29.9	55.4	29.0
Brooklyn/Curtis Bay	3,205.0	1,785.5	29.8	55.4	48.9	41.3
Canton	2,519.5	923.5	3.2	22.1	24.3	21.3
Cedonia/Frankford	4,022.5	2,250.0	73.4	40.7	49.8	34.9
Cherry Hill	1,830.5	964.5	95.9	8.5	81.6	57.9
Chinquapin Park/Belvedere	1,460.0	894.0	69.1	46.5	47.3	29.1
Claremont/Armistead	2,685.5	1,433.0	55.9	58.7	58.4	48.3
Clifton-Berea	3,248.5	1,973.5	97.0	5.7	67.6	52.5
Cross-Country/Cheswolde	2,176.5	1,574.0	19.4	40.0	15.7	25.7
Dickeyville/Franklintown	859.5	397.0	85.7	25.0	68.0	45.7
Dorchester/Ashburton	1,768.5	1,158.0	96.7	5.9	56.2	38.3
Downtown/Seton Hill	2,129.0	640.0	46.6	62.7	64.9	48.8
Edmonson Village	2,424.5	1,339.0	97.1	5.4	62.6	33.8
Fells Point	3,793.0	1,282.0	8.7	50.6	28.8	24.2
Forest Park/Walbrook	3,586.0	2,169.0	95.4	8.5	60.0	40.6
Glen-Fallstaff	2,767.5	1,635.5	61.4	53.4	39.6	36.6
Greater Charles Village/Barclay	4,858.5	2,358.0	37.1	65.0	50.5	49.5
Greater Govans	2,572.0	1,775.5	91.5	15.7	57.7	39.6
Greater Mondawmin	2,514.5	1,815.0	96.8	5.9	64.5	37.2
Greater Roland Park/Poplar Hill	1,770.5	1,339.5	6.9	29.6	13.4	13.5
Greater Rosemont	3,547.5	2,223.5	97.1	5.3	66.8	49.1
Greenmount East	2,872.0	1,670.5	96.6	6.6	69.6	58.5
Hamilton	2,599.5	1,460.0	42.5	50.5	35.3	44.2
Harford/Echodale	1,457.0	637.5	48.5	46.3	74.6	37.0
Highlandtown	4,540.5	2,915.0	30.4	53.9	39.6	26.4
Howard Park/West Arlington	2,129.5	1,340.5	52.1	45.3	34.3	32.4
Inner Harbor/Federal Hill	2,183.5	1,490.5	54.7	9.6	53.2	31.3
Jonestown/Oldtown	3,764.5	1,469.5	46.5	34.7	28.8	23.6
Lauraville	3,848.5	1,898.0	49.9	53.6	33.2	18.5
Loch Raven	2,499.5	1,460.0	83.7	27.9	51.3	27.1
Madison/East End	1,790.5	748.0	90.5	19.8	69.7	46.1
Medfield/Hampden/Woodberry/Remington	3,825.0	1,976.0	11.0	34.1	32.4	27.8
Midtown	2,982.5	1,202.0	36.4	59.7	45.8	46.2
Midway/Coldstream	1,664.5	1,267.0	96.0	7.2	65.3	40.3
Morrell Parkk/Violetville	2,478.0	1,382.5	12.4	32.9	29.8	33.7
Mount Washington/Coldspring	1,703.5	1,210.0	22.0	45.0	18.8	21.3
North Baltimore/Guilford/Homeland	3,386.5	2,086.0	12.2	38.1	14.8	23.0
Northwood	2,344.0	1,425.0	88.8	20.4	49.0	44.5
Orangeville/E. Highlandtown	3,531.0	1,955.5	50.5	17.6	80.6	44.7
Patterson Park North & East	2,023.5	1,040.0	31.5	55.6	33.1	37.9
Penn North/Reservoir Hill	3,640.0	1,857.0	65.7	67.6	51.2	36.2
Perkins/Middle East	3,145.5	1,788.0	91.1	15.0	66.5	48.9
Pimlico/Arlington/Hilltop	2,464.0	1,727.0	94.3	10.7	61.2	46.9
Poppleton/The Terraces/Hollins Market	1,454.5	671.0	82.3	30.1	74.3	62.5
Sandtown-Winchester/Harlem Park	3,990.0	2,423.0	97.1	5.3	73.2	57.4
South Baltimore	2,111.0	1,002.5	2.3	14.0	20.5	22.0
Southeastern	1,579.5	726.0	29.9	63.4	55.5	49.6
Southern Park Heights	2,147.5	1,339.0	96.4	6.7	69.6	51.4
Southwest Baltimore	4,143.5	2,233.0	73.5	42.3	63.1	48.9
The Waverlies	1,792.5	980.0	79.9	30.0	57.4	41.2
Upton/Druid Heights	2,198.0	1,071.0	93.9	11.3	77.9	67.4
Washington Village/Pigtown	1,623.5	679.0	46.7	34.3	53.2	42.5
Westport/Mount Winans/Lakeland	1,164.0	552.5	67.1	53.4	56.9	38.4

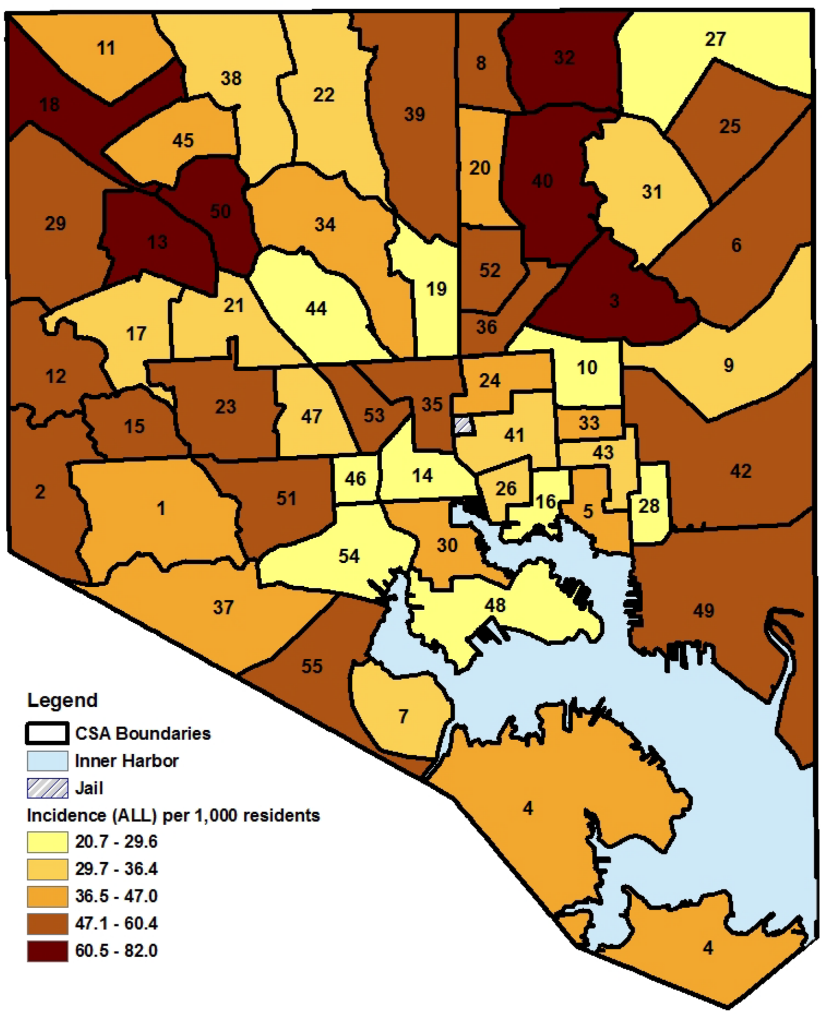
Appendix 2C Table S2.2: Characteristics of Community Statistical Areas in Baltimore City, MD (continued)

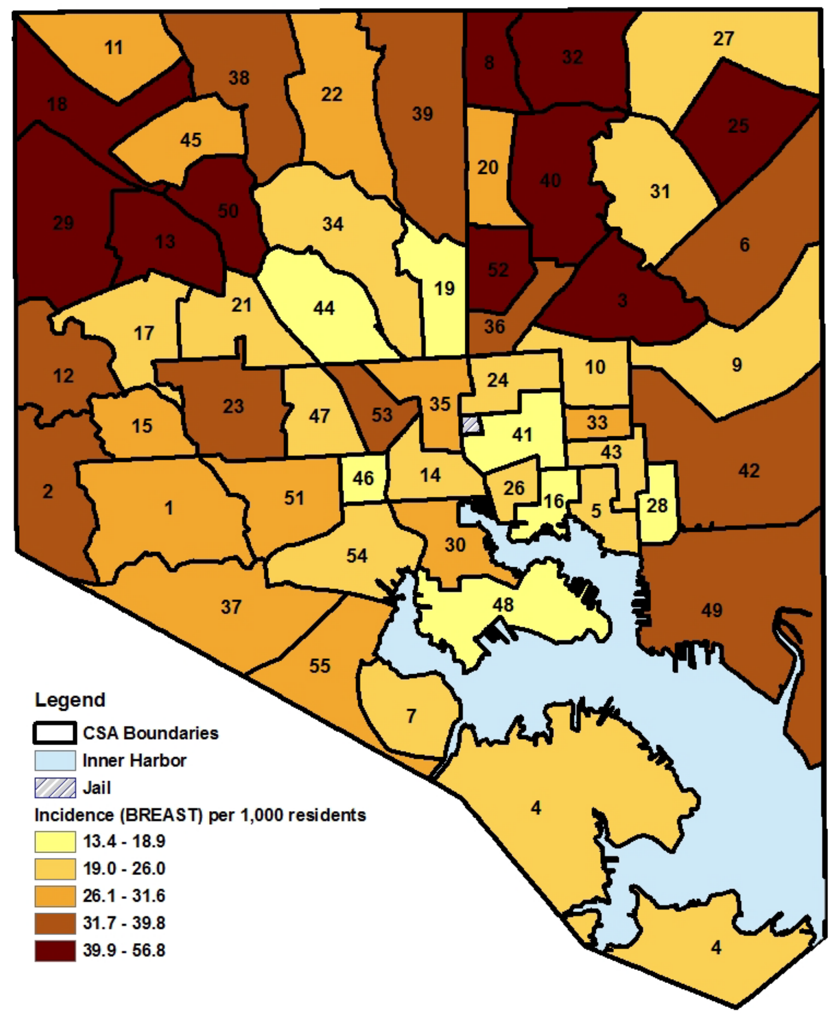
	Vacant (%)	Housing violations (%)	Crime (per 1,000)	Domestic violence (per 1,000)	Teen birth (per 1,000)	Employed (%)
Allendale/Irvington/S. Hilton	3.6	3.1	49.6	48.5	73.9	59.0
Beechfield/Ten Hills/West	0.3	1.6	39.4	37.4	59.2	68.7
Belair-Edison	1.5	2.1	55.5	50.5	77.2	62.3
Brooklyn/Curtis Bay	3.8	3.1	89.2	64.7	114.0	57.6
Canton	1.3	0.6	86.1	24.6	30.9	76.9
Cedonia/Frankford	0.6	1.3	54.5	52.2	76.1	65.9
Cherry Hill	3.2	2.1	70.3	67.7	101.2	49.0
Chinquapin Park/Belvedere	0.5	1.4	45.0	31.3	60.8	71.4
Claremont/Armistead	0.2	2.2	59.1	55.4	56.8	58.4
Clifton-Berea	23.3	5.5	60.5	49.0	98.3	45.4
Cross-Country/Cheswolde	0.2	0.5	19.3	12.5	14.4	72.9
Dickeyville/Franklintown	1.2	1.6	47.6	50.0	58.6	61.4
Dorchester/Ashburton	2.6	3.2	54.3	44.5	83.7	60.2
Downtown/Seton Hill	7.9	4.3	506.3	64.8	61.7	53.1
Edmonson Village	2.9	3.3	37.9	41.1	74.5	58.6
Fells Point	2.1	1.6	107.2	27.7	115.6	77.0
Forest Park/Walbrook	5.1	4.7	51.9	43.8	63.4	57.1
Glen-Fallstaff	0.7	1.3	62.4	33.6	44.6	64.8
Greater Charles Village/Barclay	10.0	4.1	82.0	28.3	17.9	56.1
Greater Govans	3.6	2.6	47.4	40.1	68.7	58.6
Greater Mondawmin	9.7	4.1	93.7	55.2	58.1	54.5
Greater Roland Park/Poplar Hill	0.1	0.5	35.0	14.3	0.0	76.9
Greater Rosemont	12.7	3.7	63.6	54.9	100.7	50.5
Greenmount East	31.9	7.6	66.2	48.9	97.6	43.9
Hamilton	0.3	1.4	47.4	35.6	57.7	75.0
Harford/Echodale	4.1	2.4	126.6	57.2	109.8	48.2
Highlandtown	0.2	1.1	43.4	36.1	53.1	71.8
Howard Park/West Arlington	2.1	1.3	105.9	40.3	70.8	72.3
Inner Harbor/Federal Hill	1.8	2.8	56.5	38.4	60.8	59.8
Jonestown/Oldtown	0.9	0.8	133.9	21.6	45.5	77.3
Lauraville	0.6	1.9	43.8	31.9	45.6	69.4
Loch Raven	0.1	1.5	43.0	37.4	57.4	68.9
Madison/East End	23.6	3.0	73.2	61.6	112.9	43.3
Medfield/Hampden/Woodberry/Remington	1.3	1.3	62.5	31.4	67.6	74.7
Midtown	6.7	5.0	118.5	23.0	15.5	61.7
Midway/Coldstream	15.2	5.2	61.4	50.3	78.0	52.4
Morrell Parkk/Violetville	1.2	1.8	70.9	45.2	81.8	65.2
Mount Washington/Coldspring	0.1	0.2	34.7	19.7	5.2	82.3
North Baltimore/Guilford/Homeland	0.2	0.7	33.2	13.7	2.6	45.7
Northwood	0.2	2.2	42.2	32.6	21.0	60.4
Orangeville/E. Highlandtown	37.0	2.7	93.6	50.3	79.4	43.1
Patterson Park North & East	1.6	1.9	101.5	48.5	107.3	63.7
Penn North/Reservoir Hill	8.1	2.1	80.0	48.8	80.6	61.4
Perkins/Middle East	17.5	6.8	77.0	50.5	87.0	52.0
Pimlico/Arlington/Hilltop	10.0	3.0	58.2	49.1	85.0	52.3
Poppleton/The Terraces/Hollins Market	15.8	6.6	89.4	58.7	82.6	46.7
Sandtown-Winchester/Harlem Park	27.3	6.0	71.6	60.8	92.3	47.1
South Baltimore	1.5	0.6	58.4	20.6	19.8	75.1
Southeastern	0.5	1.2	69.3	57.3	87.5	57.2
Southern Park Heights	14.1	5.2	54.4	50.4	98.6	47.6
Southwest Baltimore	21.8	4.8	79.7	65.0	107.2	48.5
The Waverlies	3.8	2.9	75.2	51.9	81.5	62.6
Upton/Druid Heights	28.6	7.0	90.4	59.7	109.1	42.1
Washington Village/Pigtown	10.8	3.0	135.3	55.0	99.1	57.1
Westport/Mount Winans/Lakeland	6.6	2.4	89.0	46.5	98.1	58.7

Appendix 2C Table S2.2: Characteristics of Community Statistical Areas in Baltimore City, MD (continued)

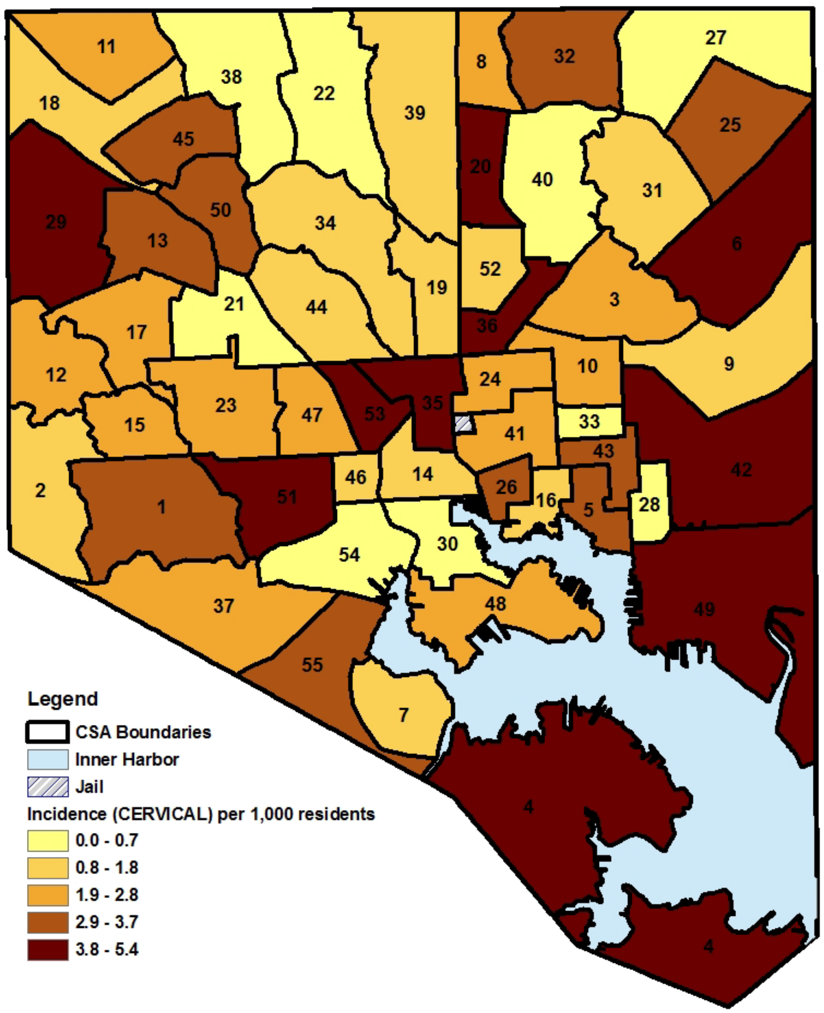
	Businesses (n)	Voted (%)	Dirty streets (per 1,000)	Tree coverage (%)	Neighborhood associations (n)
Allendale/Irvington/S. Hilton	271.3	36.7	32.0	29.4	11.5
Beechfield/Ten Hills/West	176.3	44.2	11.4	44.3	5.5
Belair-Edison	229.5	40.5	83.8	16.4	4.5
Brooklyn/Curtis Bay	472.0	21.8	70.6	15.1	6.0
Canton	350.0	42.0	38.8	6.5	7.0
Cedonia/Frankford	421.0	37.9	26.7	25.4	2.3
Cherry Hill	123.8	34.6	8.9	16.1	3.5
Chinquapin Park/Belvedere	143.3	45.9	23.5	34.1	6.0
Claremont/Armistead	204.3	28.7	10.0	26.6	1.8
Clifton-Berea	181.5	36.4	113.7	5.6	11.3
Cross-Country/Cheswolde	204.3	54.9	5.8	41.2	2.5
Dickeyville/Franklintown	46.8	31.7	2.0	68.7	4.5
Dorchester/Ashburton	168.0	45.3	27.7	24.0	9.5
Downtown/Seton Hill	3,608.5	24.0	31.7	4.1	6.0
Edmonson Village	60.5	45.0	28.6	49.1	6.3
Fells Point	491.8	35.2	43.4	4.0	9.5
Forest Park/Walbrook	151.0	37.9	37.6	47.0	9.3
Glen-Fallstaff	708.5	43.4	23.0	22.9	9.0
Greater Charles Village/Barclay	942.5	29.0	54.2	17.4	11.3
Greater Govans	143.8	40.9	40.1	27.0	12.3
Greater Mondawmin	253.5	40.4	84.9	17.5	10.3
Greater Roland Park/Poplar Hill	372.5	59.7	12.0	50.8	8.8
Greater Rosemont	324.0	37.2	71.1	20.9	29.3
Greenmount East	189.5	31.4	112.8	9.4	9.8
Hamilton	282.0	43.2	22.6	28.9	4.5
Harford/Echodale	585.5	37.9	28.2	4.0	12.3
Highlandtown	313.3	36.4	13.3	25.6	3.8
Howard Park/West Arlington	413.8	43.5	95.2	1.8	4.0
Inner Harbor/Federal Hill	187.3	44.5	24.8	35.9	6.5
Jonestown/Oldtown	903.5	31.0	38.8	5.0	12.5
Lauraville	244.0	44.6	21.8	34.5	8.5
Loch Raven	251.0	43.1	30.4	29.7	6.3
Madison/East End	208.8	26.9	311.4	3.5	8.3
Medfield/Hampden/Woodberry/Remington	913.5	40.0	34.6	27.7	9.8
Midtown	1,188.0	34.2	36.7	8.2	9.3
Midway/Coldstream	234.3	34.1	107.6	7.4	3.0
Morrell Parkk/Violetville	520.5	29.6	20.6	21.5	5.8
Mount Washington/Coldspring	216.0	63.7	8.6	59.8	4.5
North Baltimore/Guilford/Homeland	450.8	46.8	13.0	43.4	15.3
Northwood	136.8	44.0	22.5	25.5	12.0
Orangeville/E. Highlandtown	525.5	27.5	51.0	5.9	7.5
Patterson Park North & East	593.5	29.6	84.1	9.9	3.5
Penn North/Reservoir Hill	220.0	33.3	207.2	3.1	8.3
Perkins/Middle East	223.3	30.2	66.7	37.4	16.8
Pimlico/Arlington/Hilltop	428.3	34.0	54.0	18.3	16.8
Poppleton/The Terraces/Hollins Market	155.0	31.0	90.5	8.5	9.5
Sandtown-Winchester/Harlem Park	255.5	30.0	142.0	10.2	17.5
South Baltimore	232.3	41.8	30.5	3.6	3.8
Southeastern	420.3	25.7	22.0	3.3	6.8
Southern Park Heights	196.0	32.1	60.5	23.0	19.0
Southwest Baltimore	481.8	25.7	155.2	10.1	9.5
The Waverlies	197.3	42.5	59.4	14.3	3.5
Upton/Druid Heights	293.5	30.7	65.9	8.7	16.5
Washington Village/Pigtown	354.5	30.2	133.6	8.9	6.8
Westport/Mount Winans/Lakeland	256.0	29.3	40.2	18.5	3.8

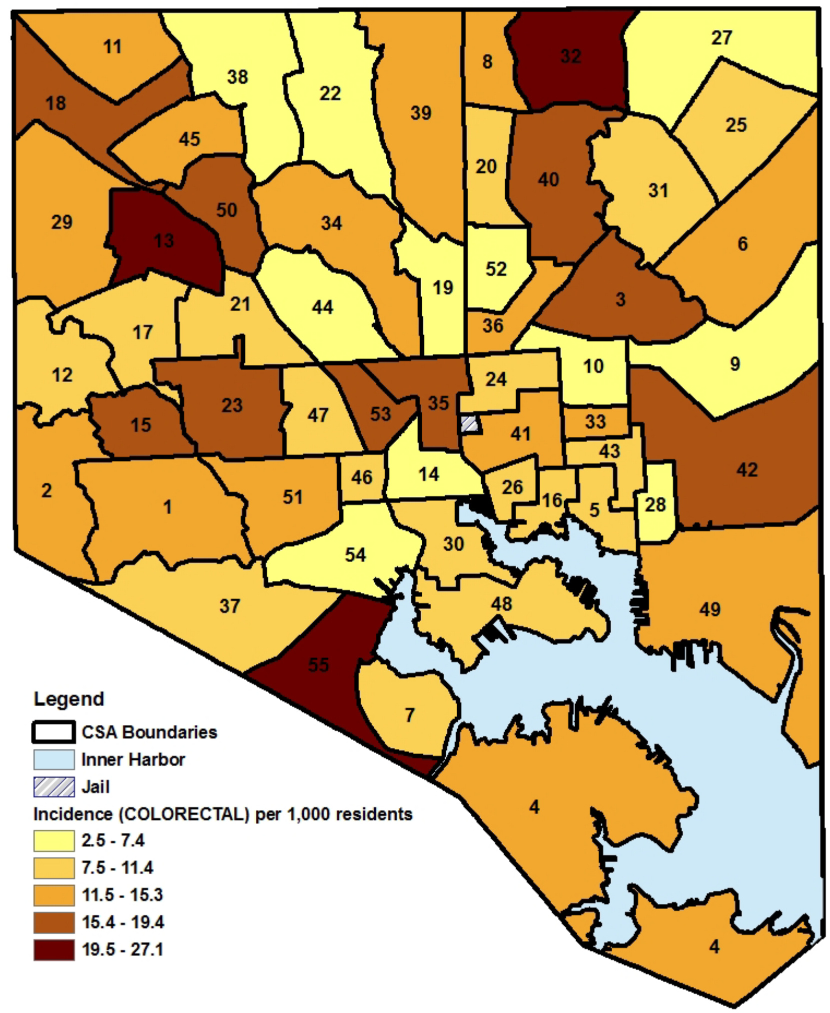
Appendix 2D Figure S2.2a-b: CSA distribution shaded by quintile of female cancer incidence (all and breast) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years



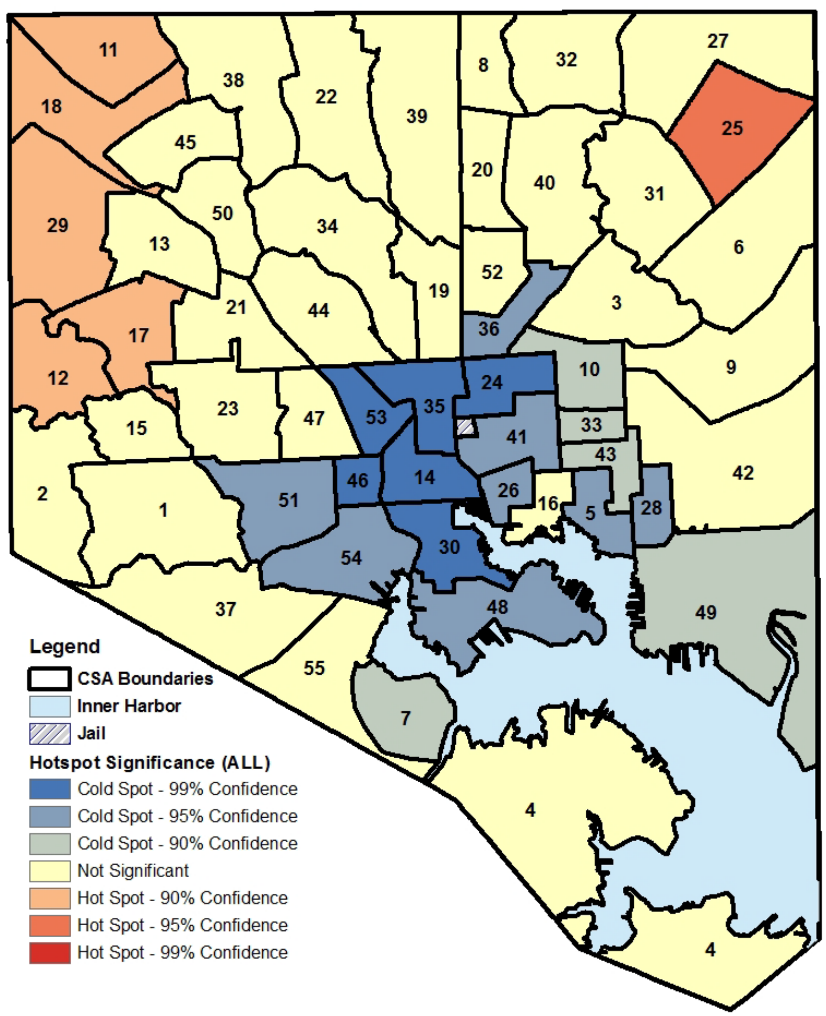


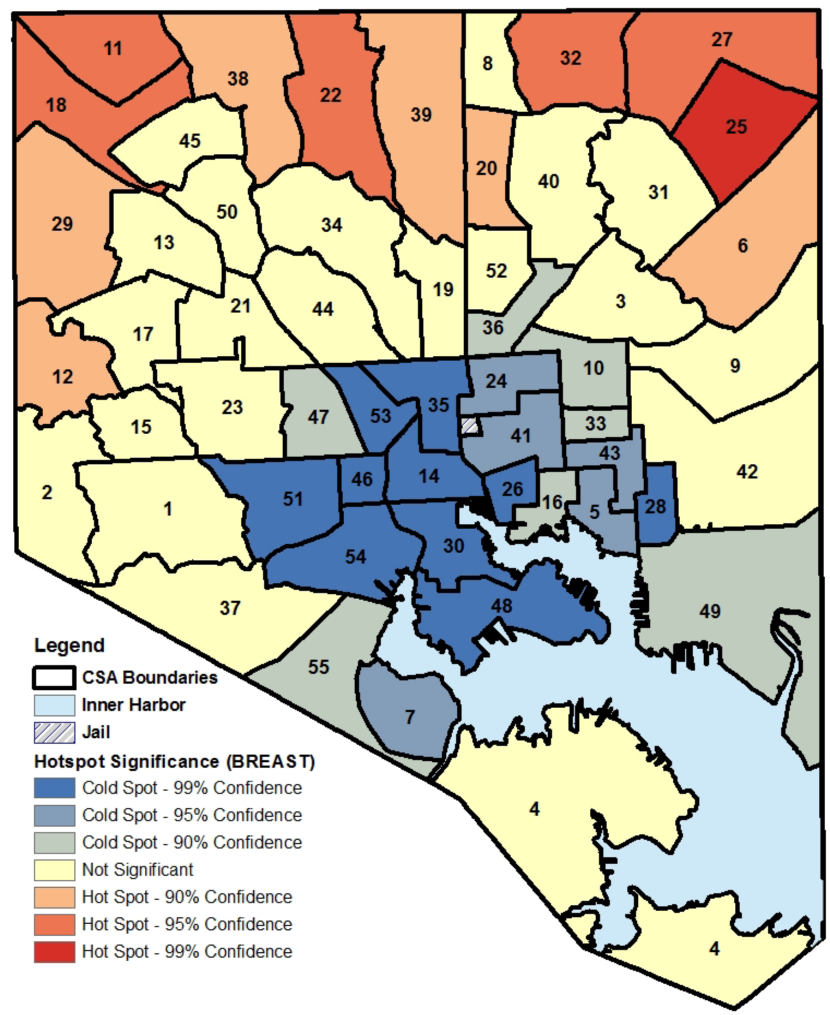
Appendix 2D Figure S2.2c-d: CSA distribution shaded by quintile of female cancer incidence (cervical and colorectal) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years



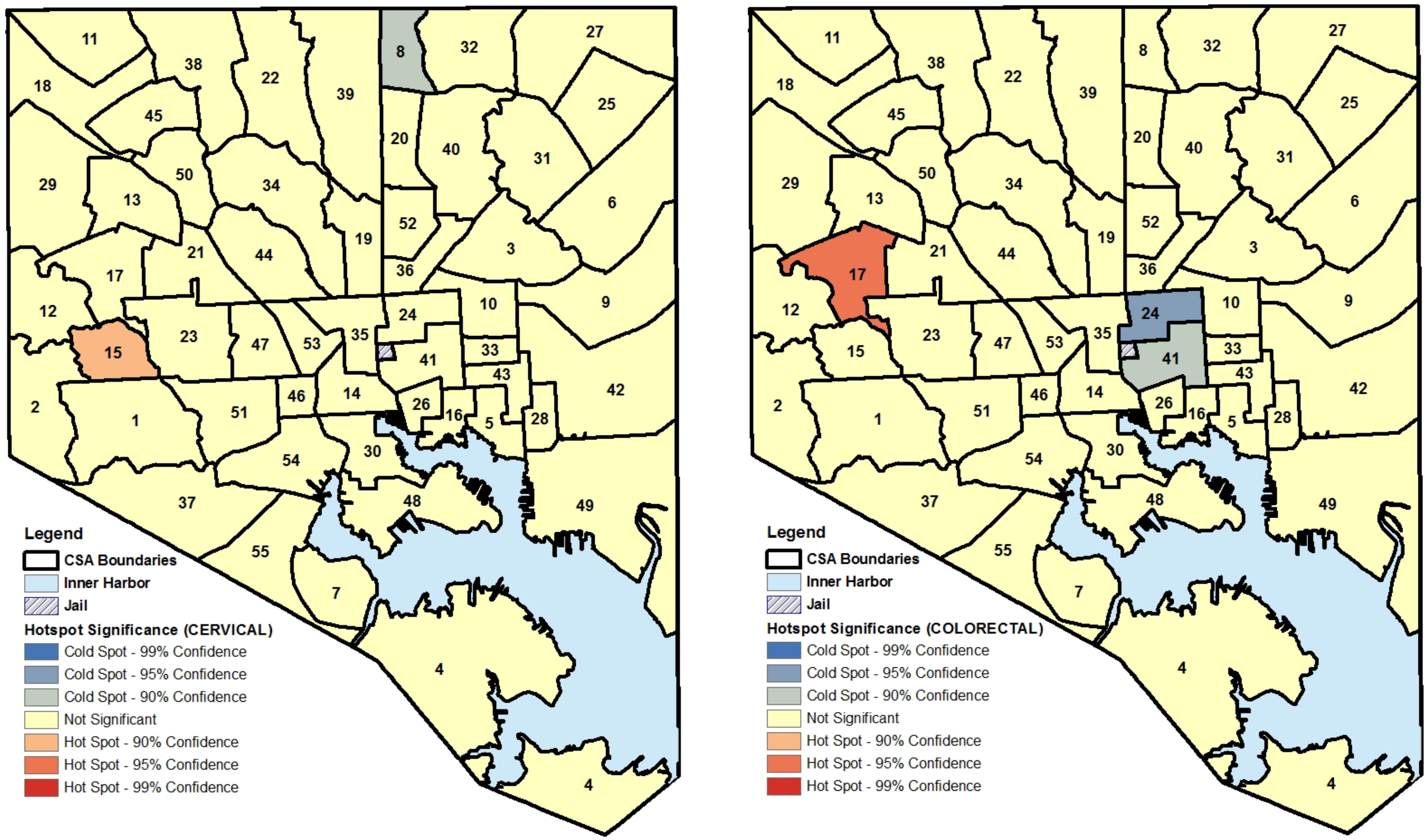


Appendix 2D Figure S2.3a-b: Hot spot analysis and statistical significance of cancer incidence (all and breast) per 1,000 female residents aged 50 to 74 years by CSA





Appendix 2D Figure S2.3c-d: Hot spot analysis and statistical significance of cancer incidence (cervical and colorectal) per 1,000 female residents aged 50 to 74 years by CSA

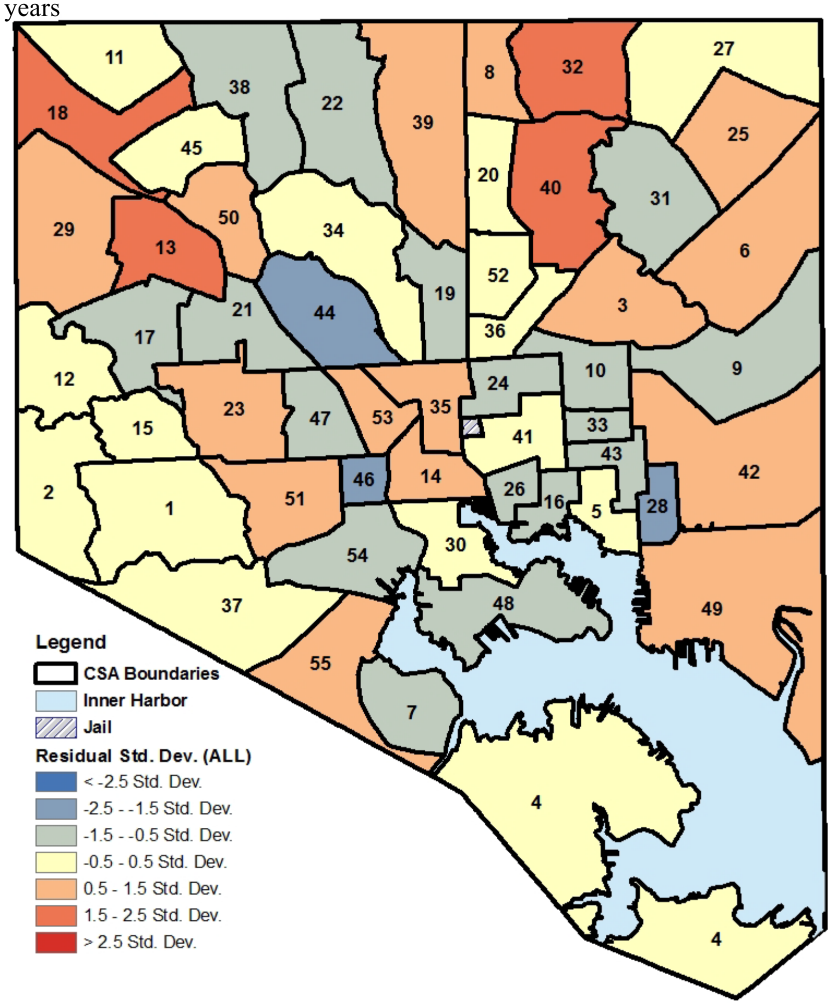


Appendix 2D Table S2.3: Ordinary Least Squares regression models for cancer incidence among females 50-74 years by site and candidate neighborhood-level covariates

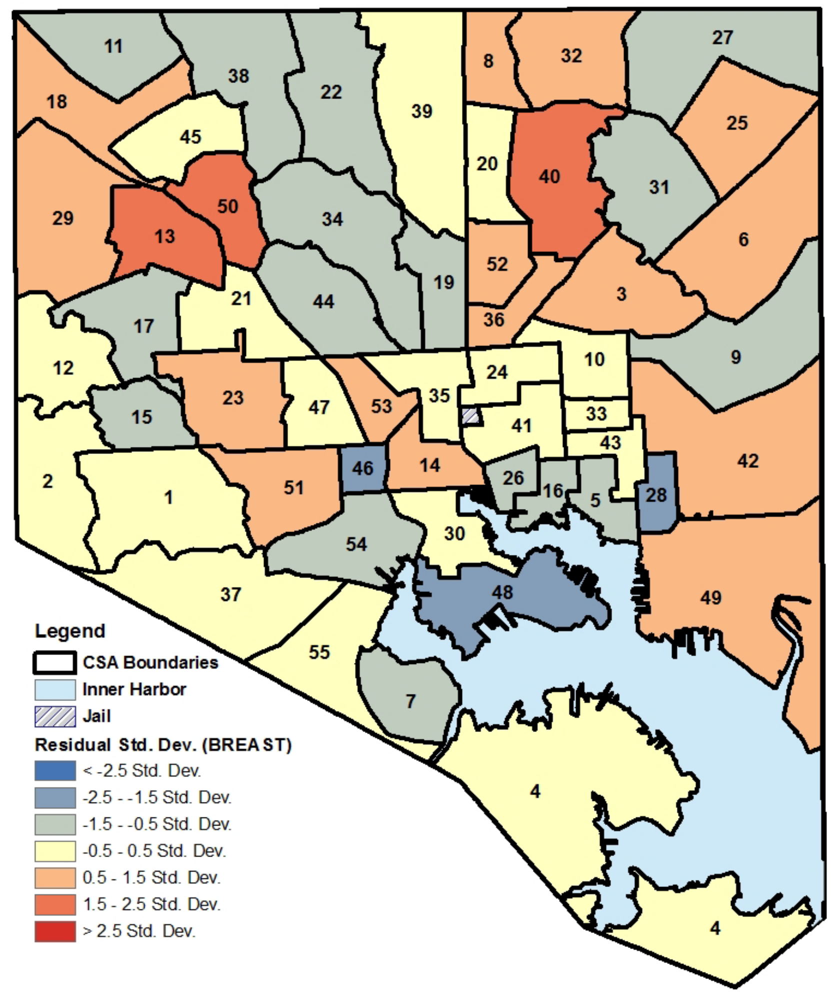
	All Cancers			Breast			Cervical			Colorectal	
	Coefficient	p-value		Coefficient	p-value		Coefficient	p-value		Coefficient	p-value
Unadjusted											
Females 50-74 yrs	-0.004	0.329		-0.002	0.423		0.00001	0.986		-0.001	0.255
% AA	0.146	0.021*		0.088	0.054		0.008	0.211		0.051	0.023*
Racial Diversity	-0.103	0.313		-0.060	0.413		-0.003	0.739		-0.040	0.269
Household income <25K	-0.008	0.964		-0.071	0.556		0.030	0.054		0.033	0.575
Female headed	0.049	0.665		-0.026	0.750		0.019	0.065		0.055	0.161
Vacants	-0.315	0.143		-0.321	0.035*		0.015	0.457		-0.009	0.906
Housing violations	-0.853	0.432		-1.104	0.154		0.180	0.075		0.072	0.852
Crime	-0.063	0.041*		-0.046	0.038*		-0.001	0.678		0.036	0.464
Domestic violence	-0.009	0.952		-0.069	0.497		0.024	0.070		0.041	0.070
Teen births	0.021	0.748		-0.034	0.464		0.014	0.019*		-0.080	0.227
Employed	-0.006	0.975		0.104	0.444		-0.029	0.095		-0.002	0.239
Businesses	-0.006	0.136		-0.004	0.145		0.000	0.765		-0.036	0.664
Voted	0.311	0.187		0.390	0.019*		-0.043	0.049*		-0.005	0.689
Dirty streets	-0.047	0.195		-0.041	0.111		-0.001	0.878		-0.004	0.929
Tree coverage	0.175	0.164		0.191	0.031*		-0.013	0.285		0.110	0.430
Neighborhood associations	0.175	0.658		0.044	0.877		0.021	0.570		-0.001	0.255
Adjusted											
Females 50-74 yrs	-0.007	0.064	Females 50-74 yrs	-0.003	0.228	Females 50-74 years	0.0001		Females 50-74 yrs	-0.002	0.178
% AA	0.138	0.022*	Vacants	-0.160	0.365	Teen births	0.0112		% AA	0.053	0.018*
Crime	-0.072	0.019*	Crime	-0.037	0.139	Voted	-0.0174				
			Voted	0.126	0.562						
			Tree coverage	0.063	0.568						
R-squared	0.166			0.100			0.056			0.092	

* Statistically significant

Appendix 2D Figure S2.4a-b: Spatial output of final models for ordinary least squares regression for cancer incidence (all and breast) of females aged 50 to 74 years

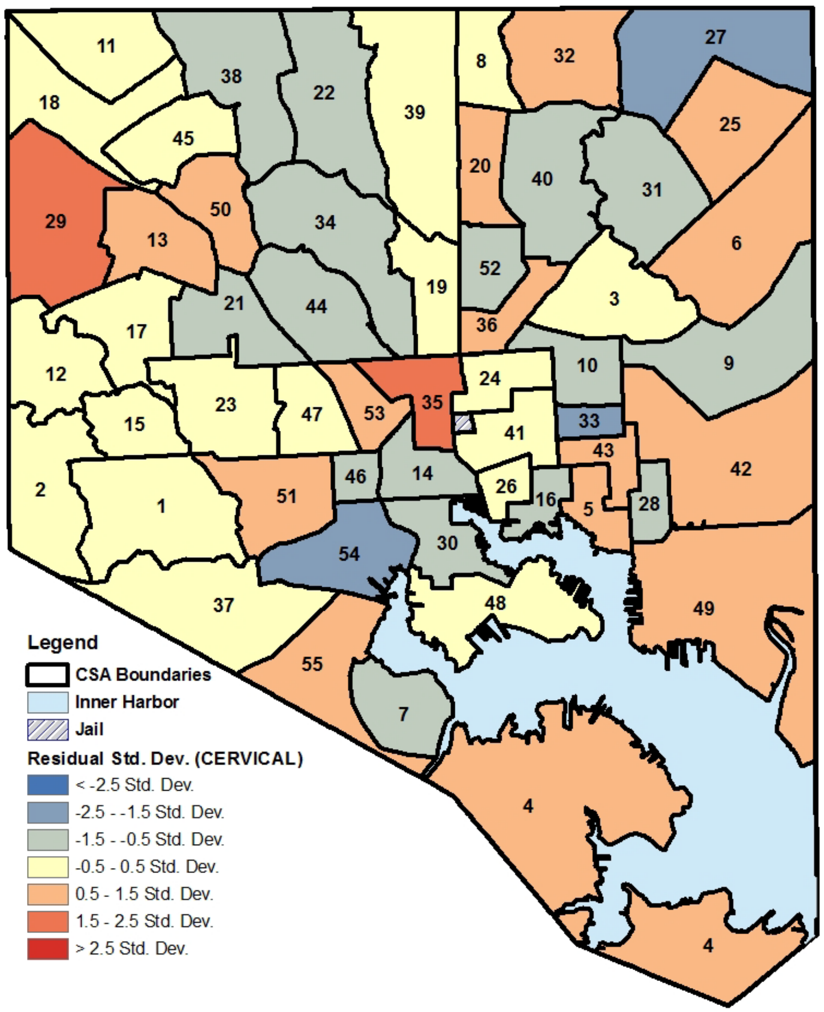


[a]Overall cancer model: Females (50-74 years), % African-American, crime coverage

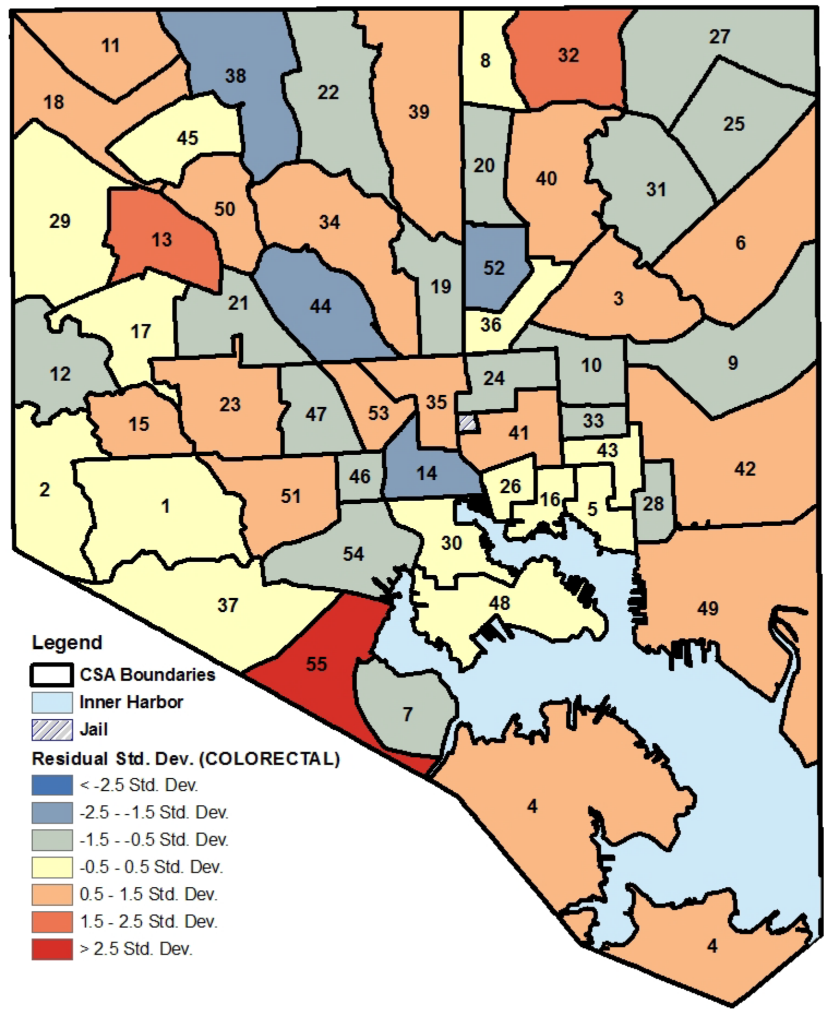


[b]Breast cancer model: Females (50-74 years), vacants, crime, voted, tree

Appendix 2D Figure S2.4c-d: Spatial output of final models for ordinary least squares regression for cancer incidence (cervical and colorectal) in females aged 50 to 74 years

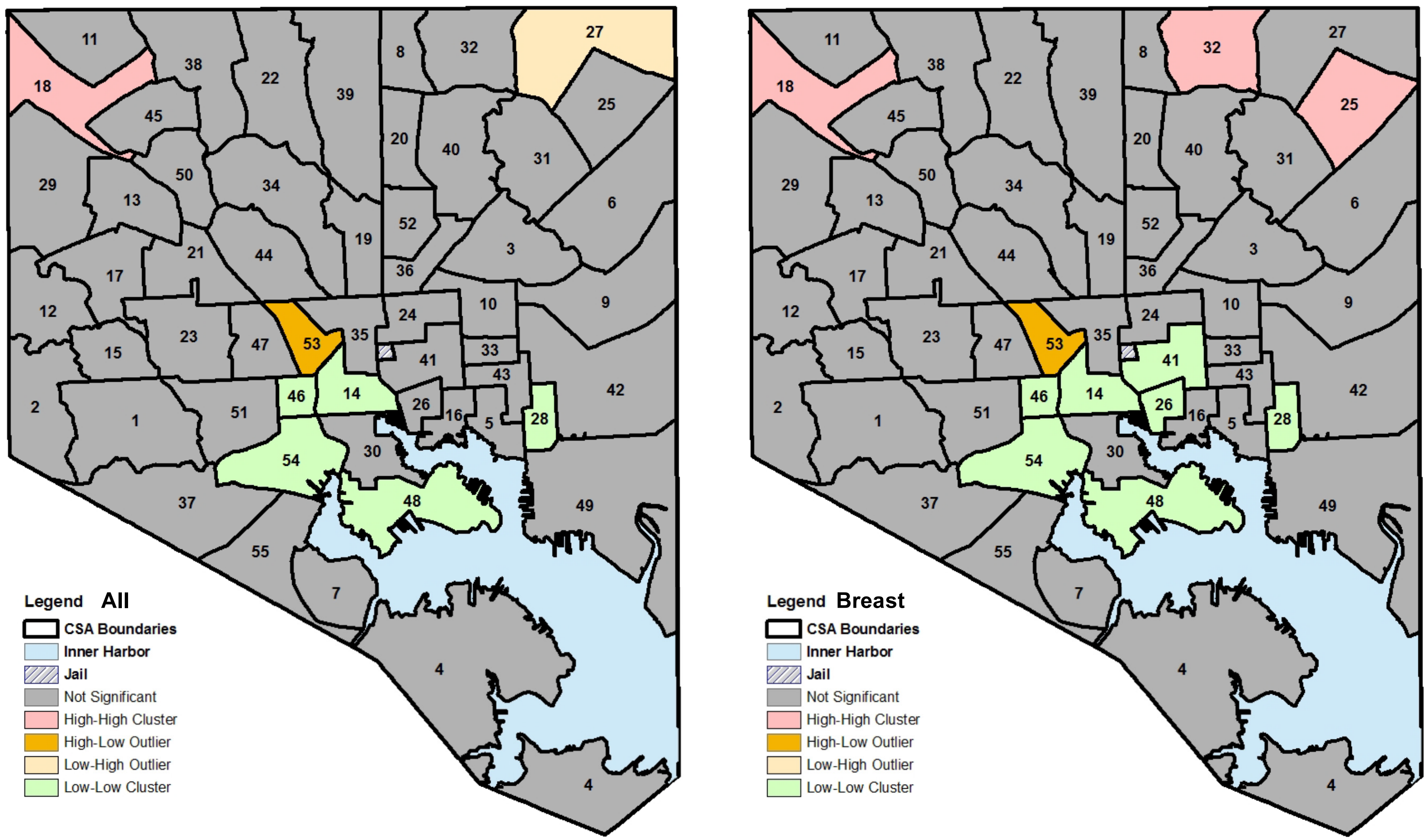


[a]Overall cancer model: Females (50-74 years), teen births, voted

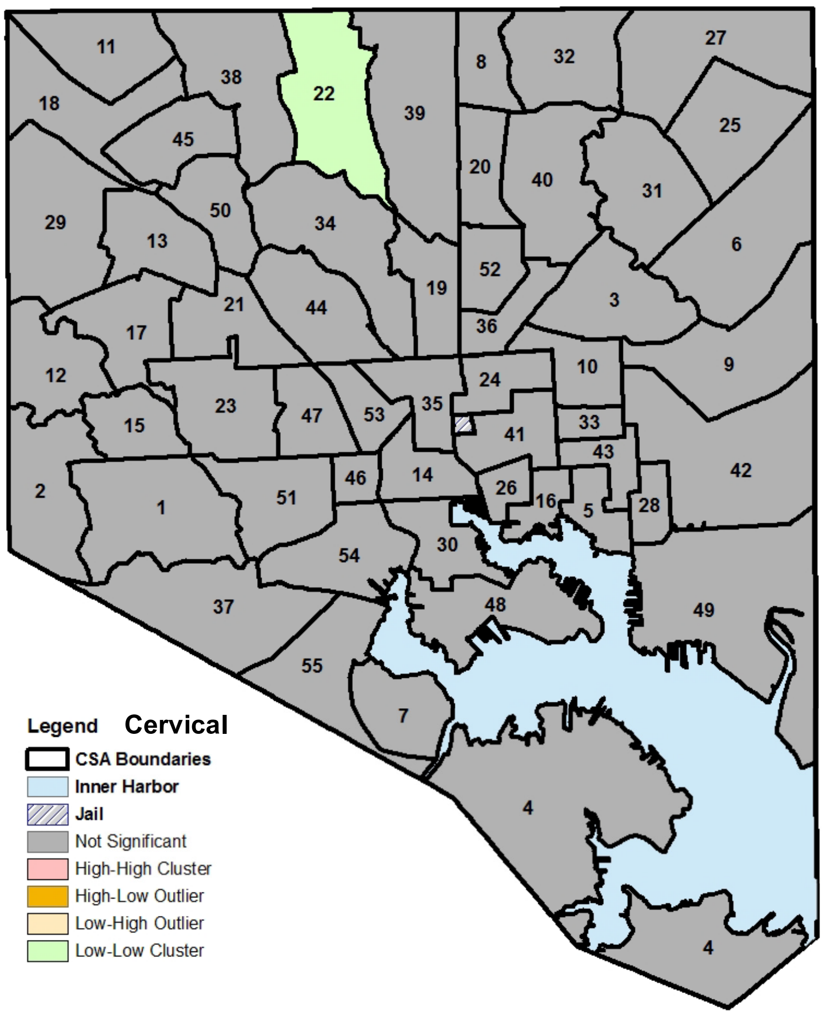


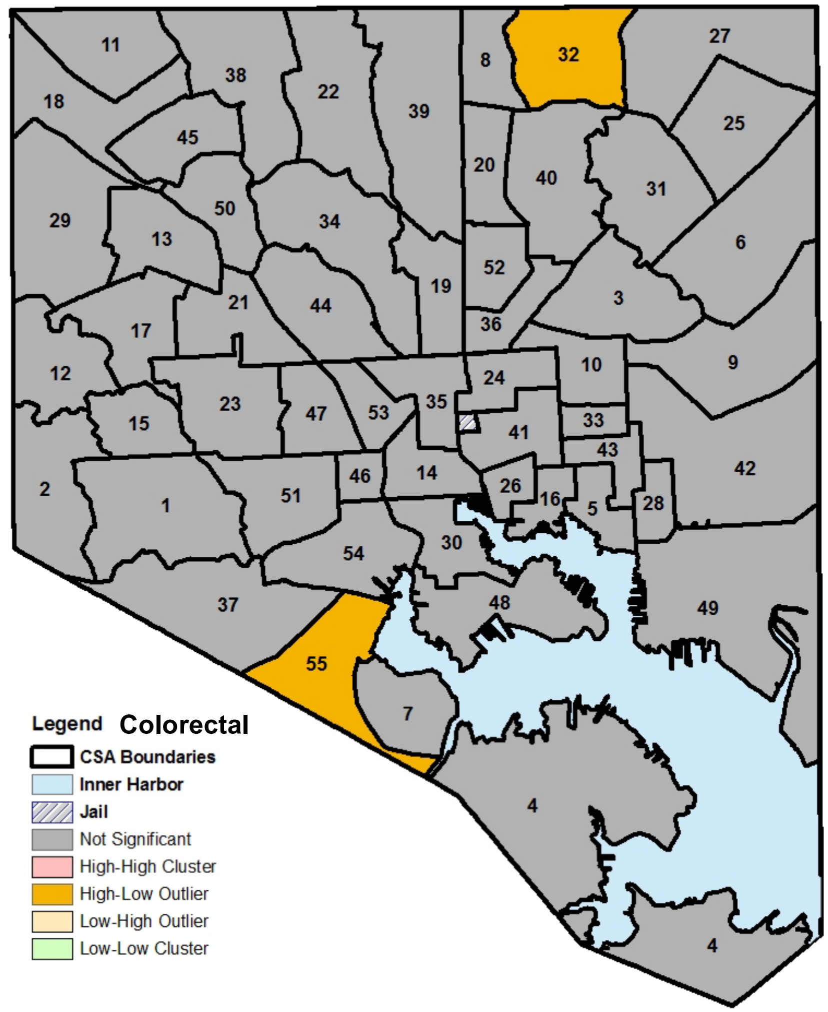
[b]Breast cancer model: Females (50-74 years), % African-American

Appendix 2D Figure S2.5a-b: Local Moran's I analysis for cancer incidence (all and breast) in females aged 50 to 74 years

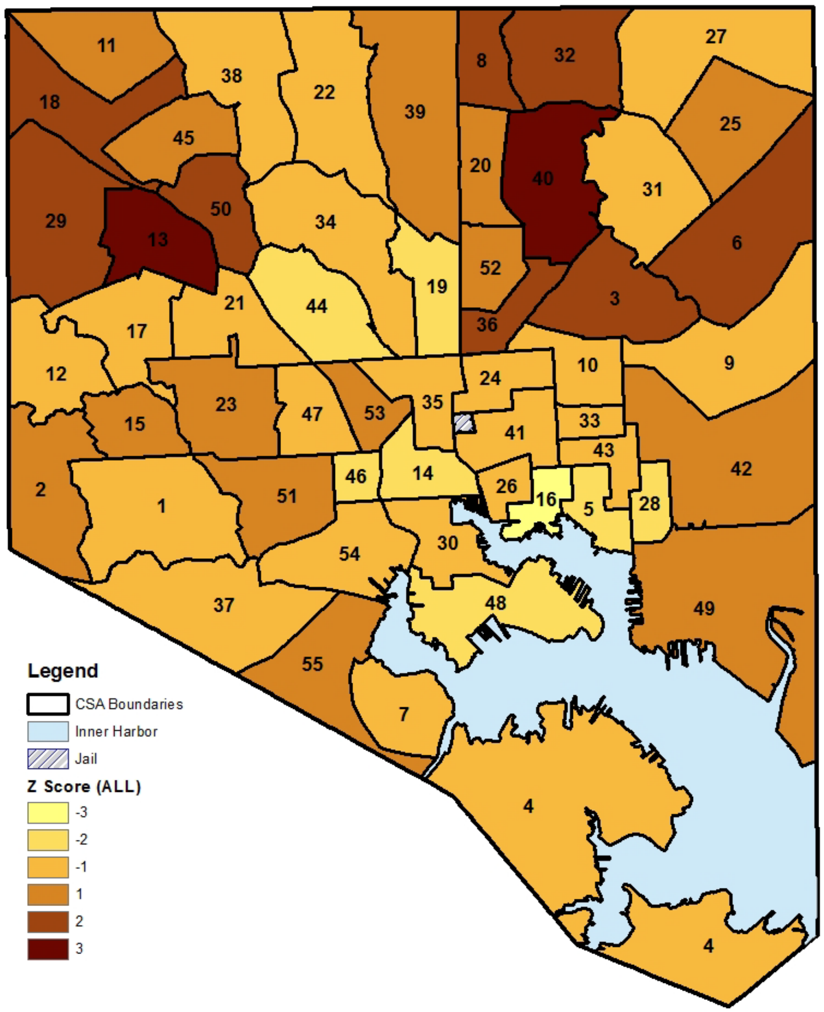


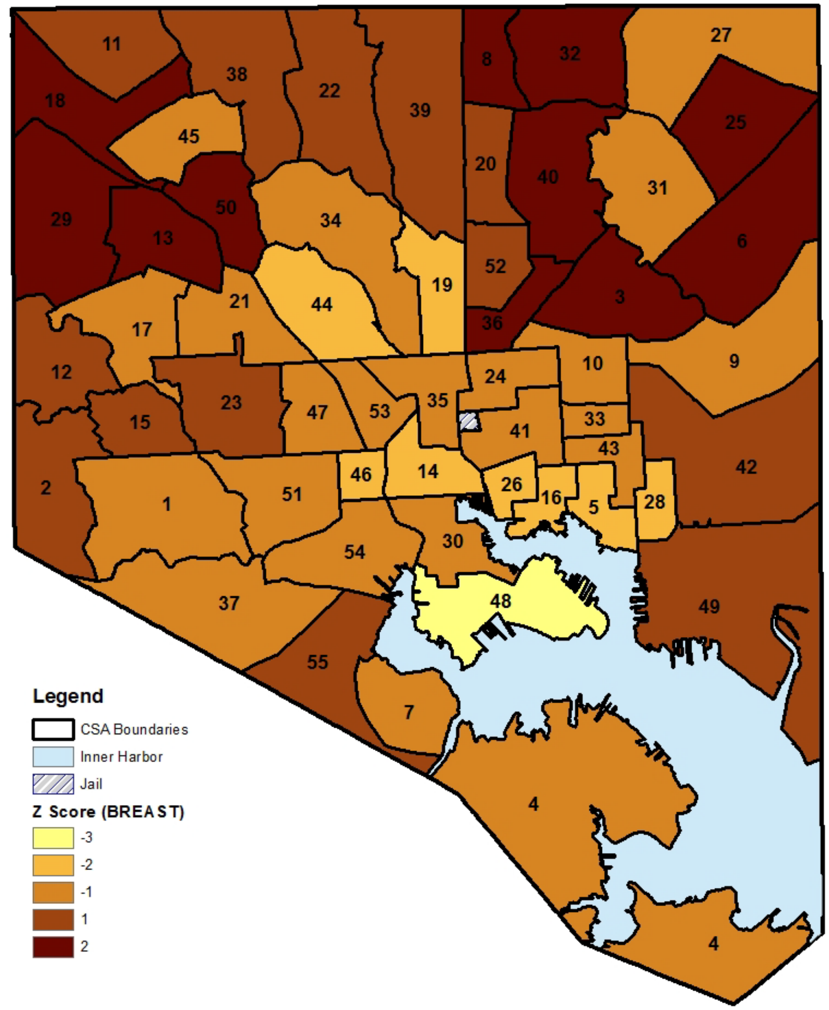
Appendix 2D Figure S2.5c-d: Local Moran's I analysis for cancer incidence (cervical and colorectal) in females aged 50 to 74 years



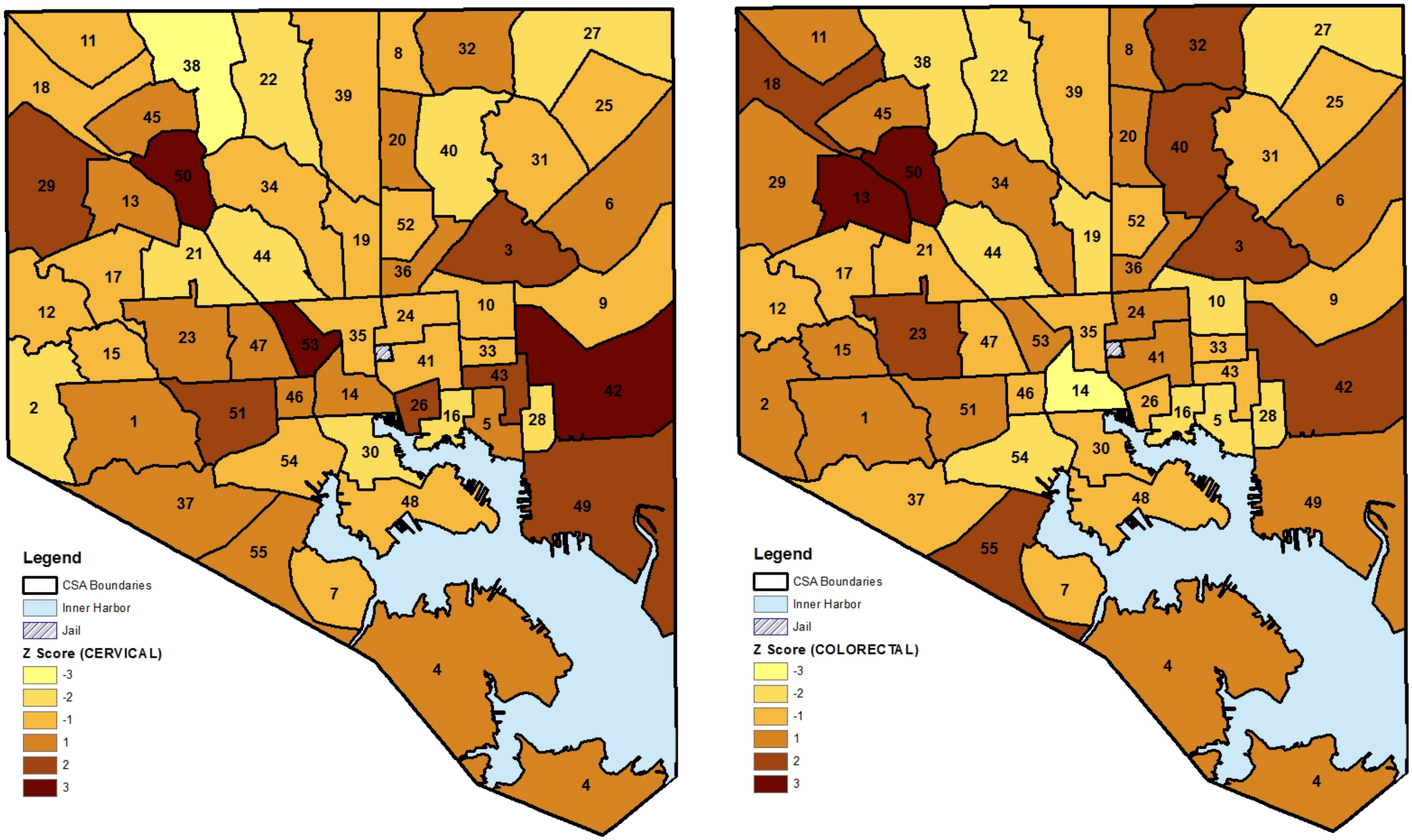


Appendix 2E Figure S2.6a-b: CSA distribution shaded by Z-score of female cancer incidence (all and breast) in Baltimore City, MD per 1,000 female residents aged 21 to 74 years

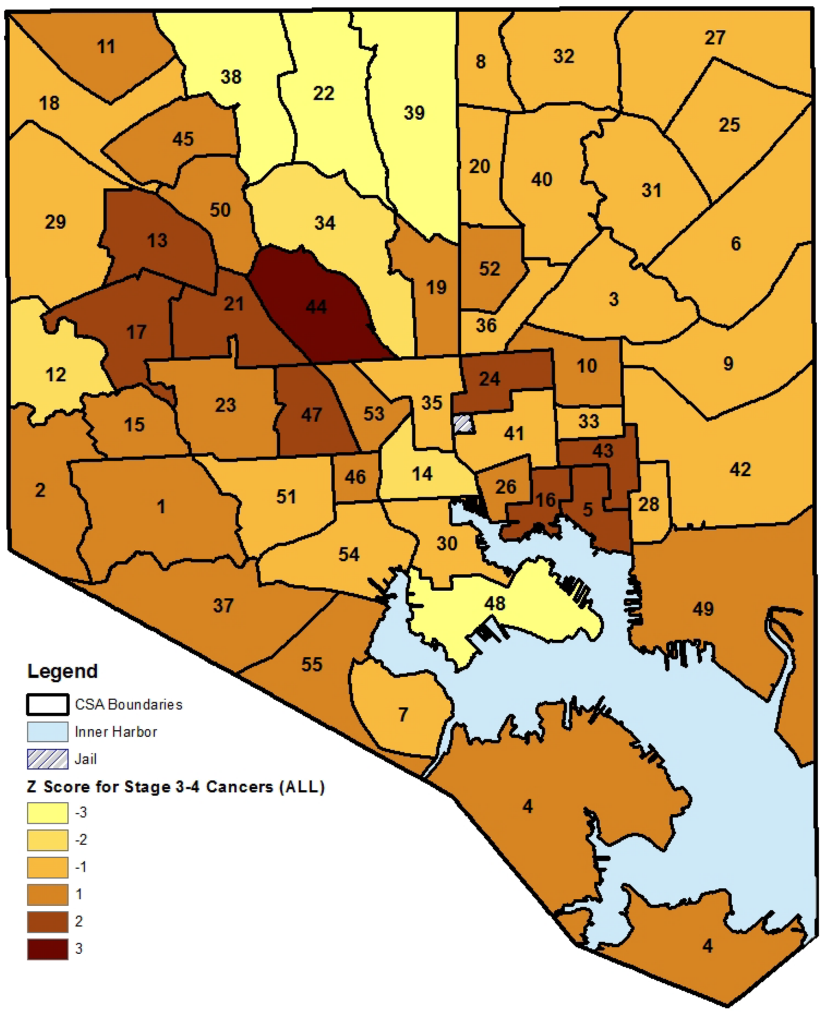


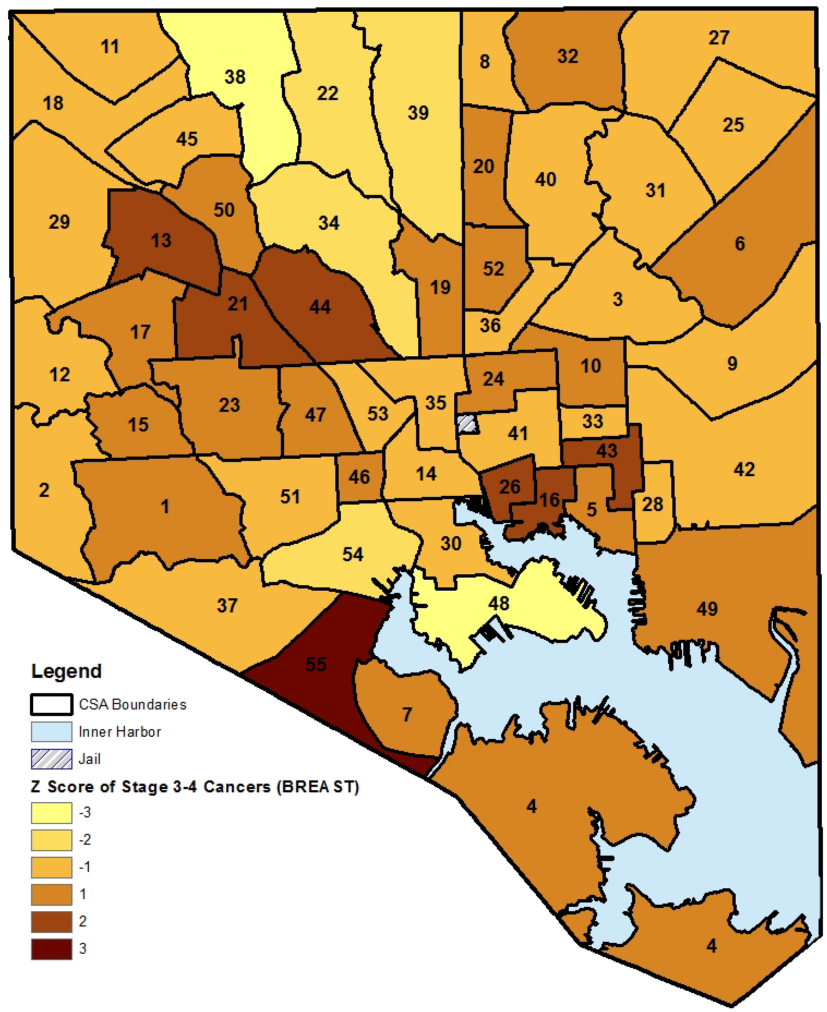


Appendix 2E Figure S2.6c-d: CSA distribution shaded by z-score of female cancer incidence (cervical and colorectal) in Baltimore City, MD per 1,000 female residents aged 21 to 74 years

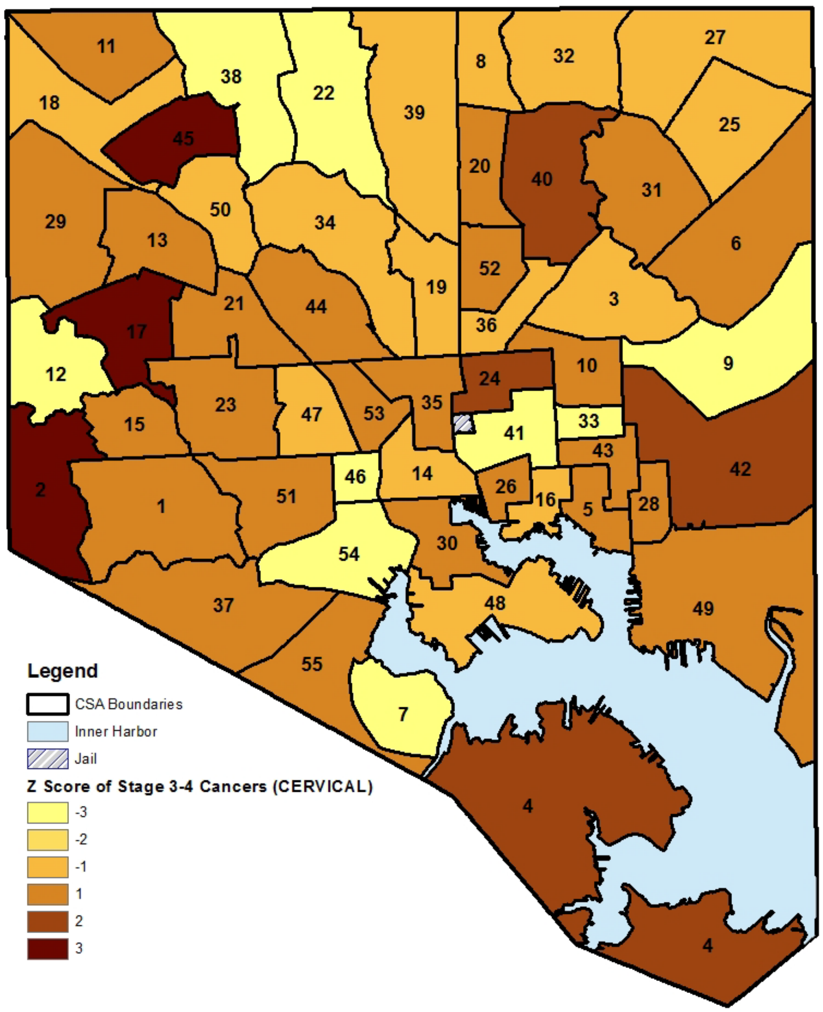


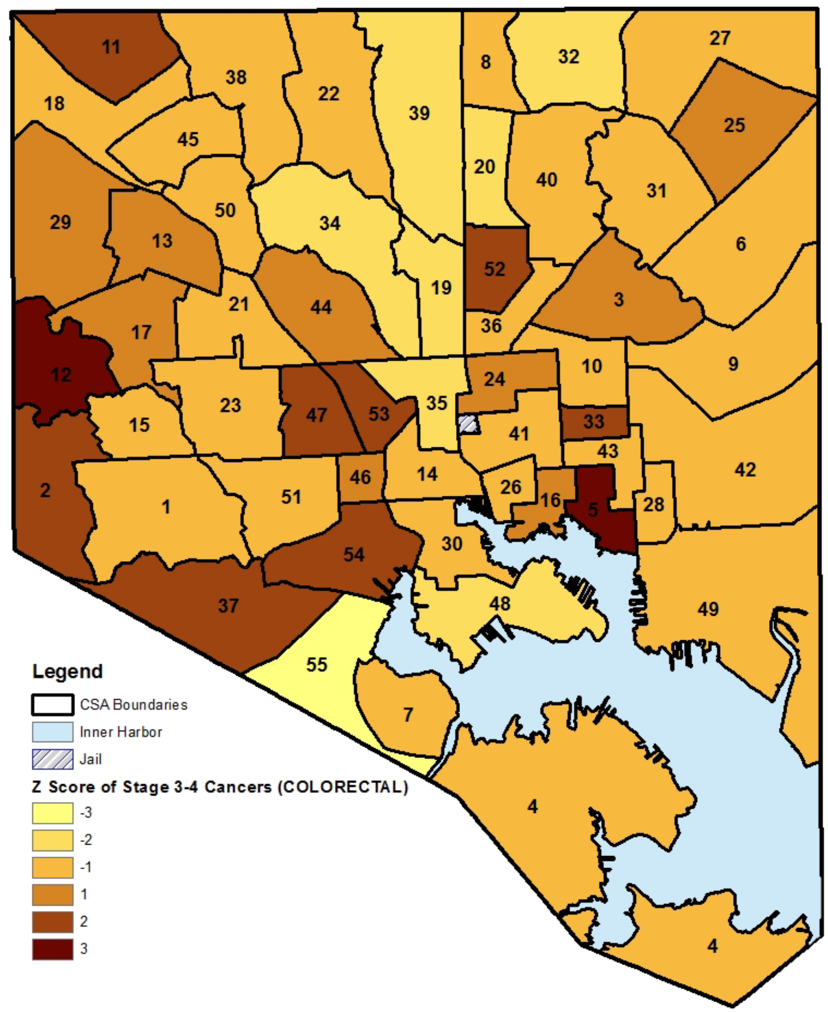
Appendix 2F Figure S2.7a-b: CSA distribution shaded by Z-Score of stage 3 and 4 cancers over total cancers diagnosed (all and breast) in Baltimore City, MD



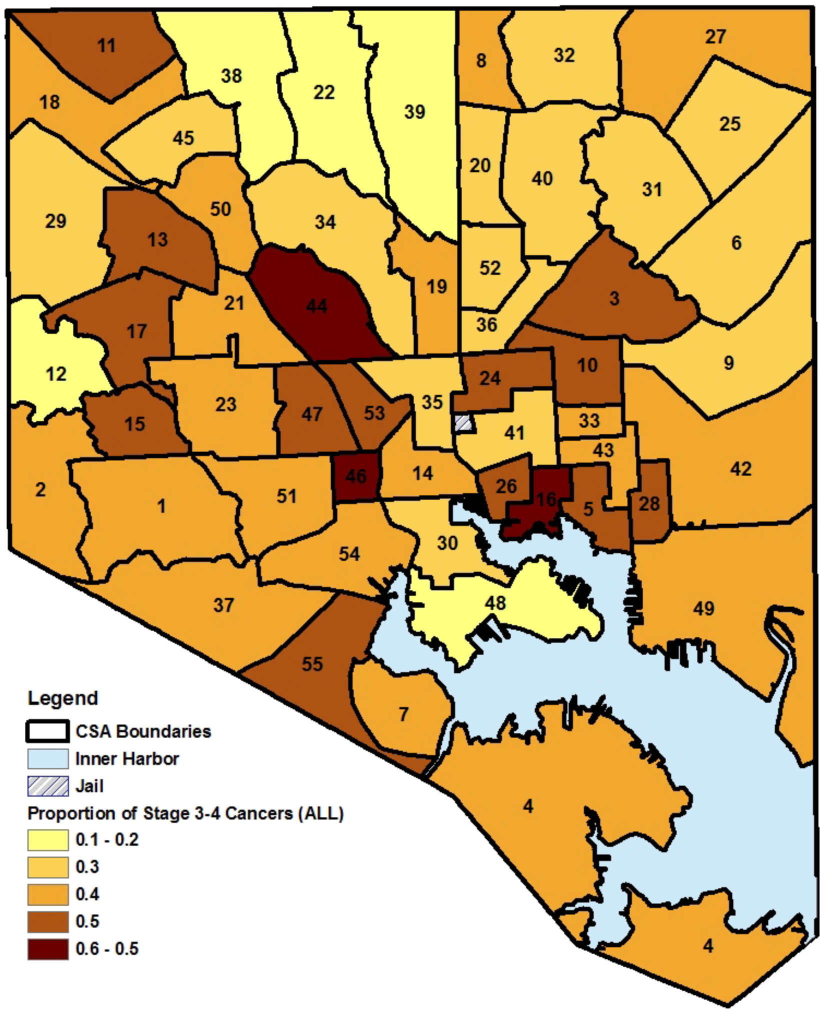


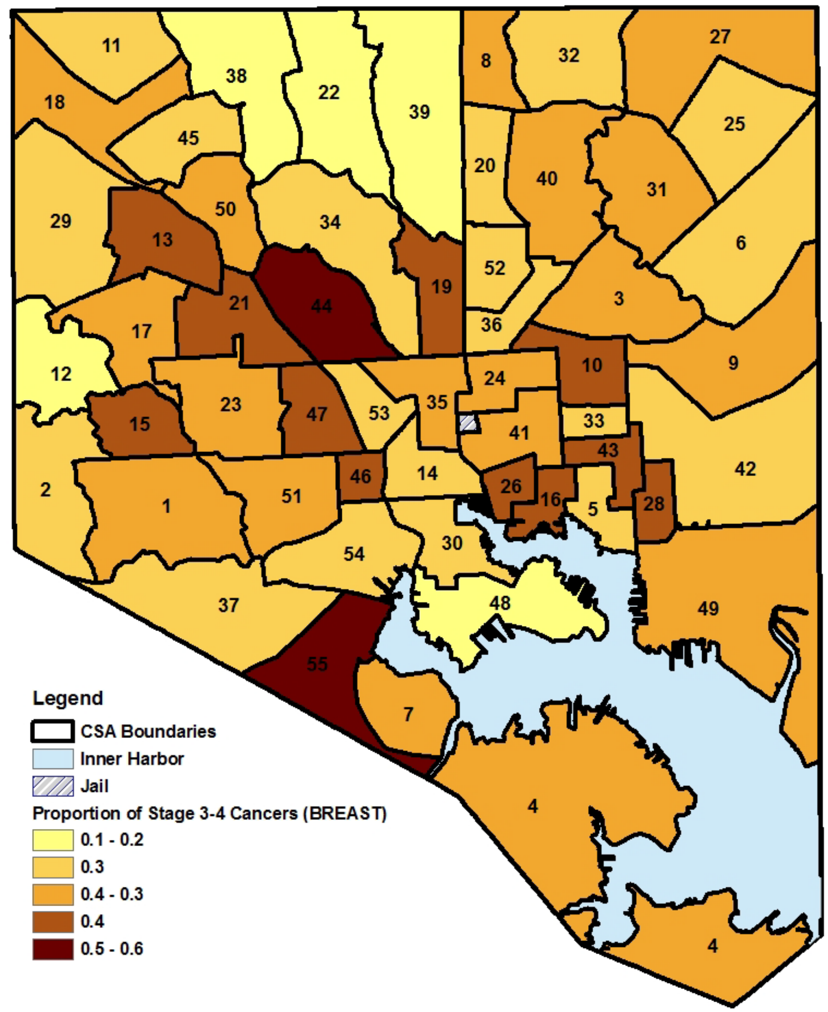
Appendix 2F Figure S2.7c-d: CSA distribution shaded by Z-score of stage 3 and 4 cancers over total cancers diagnosed (cervical and colorectal) in Baltimore City, MD



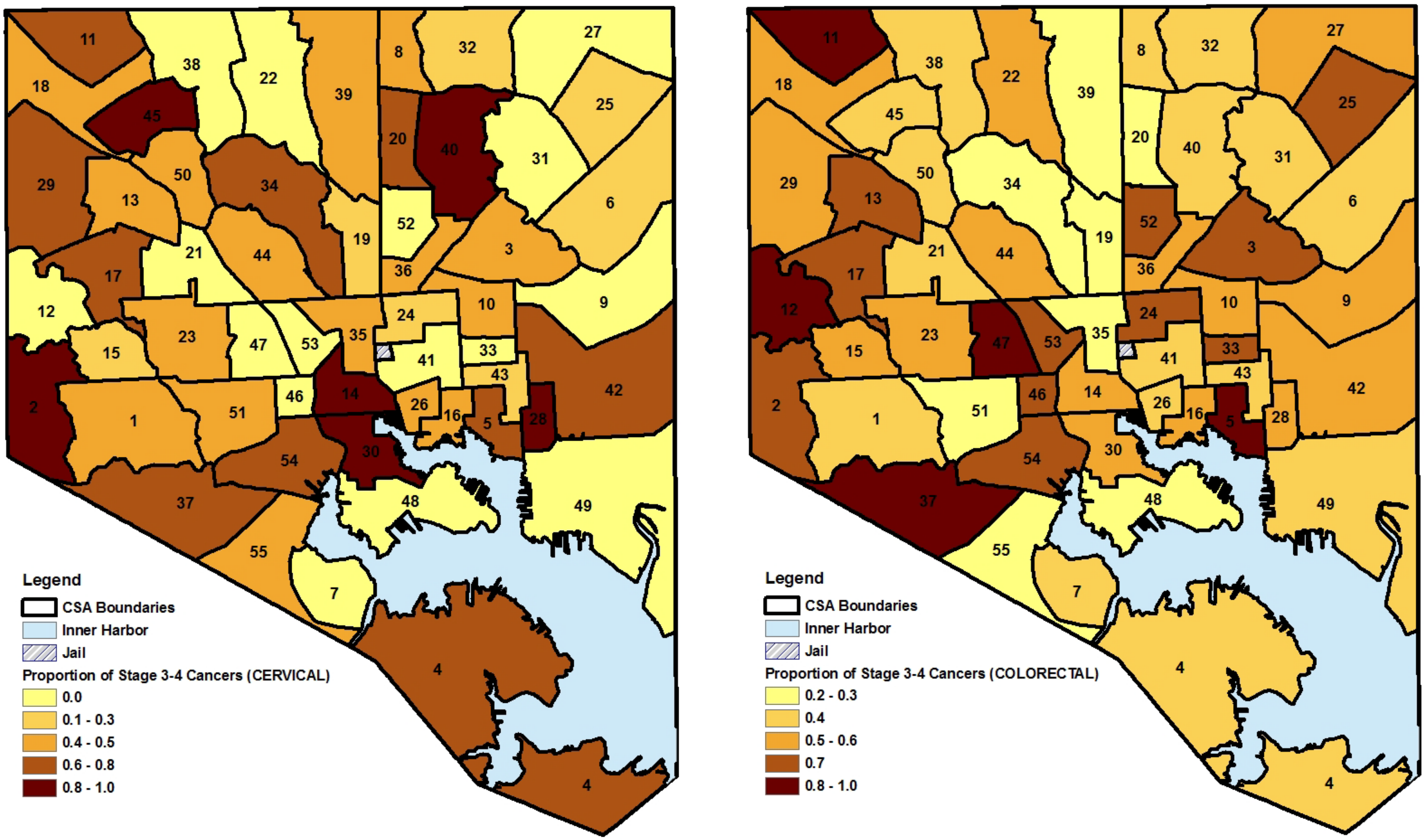


Appendix 2G Figure S2.8a-b: CSA distribution shaded by quintile of cancer stage (all and breast) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years

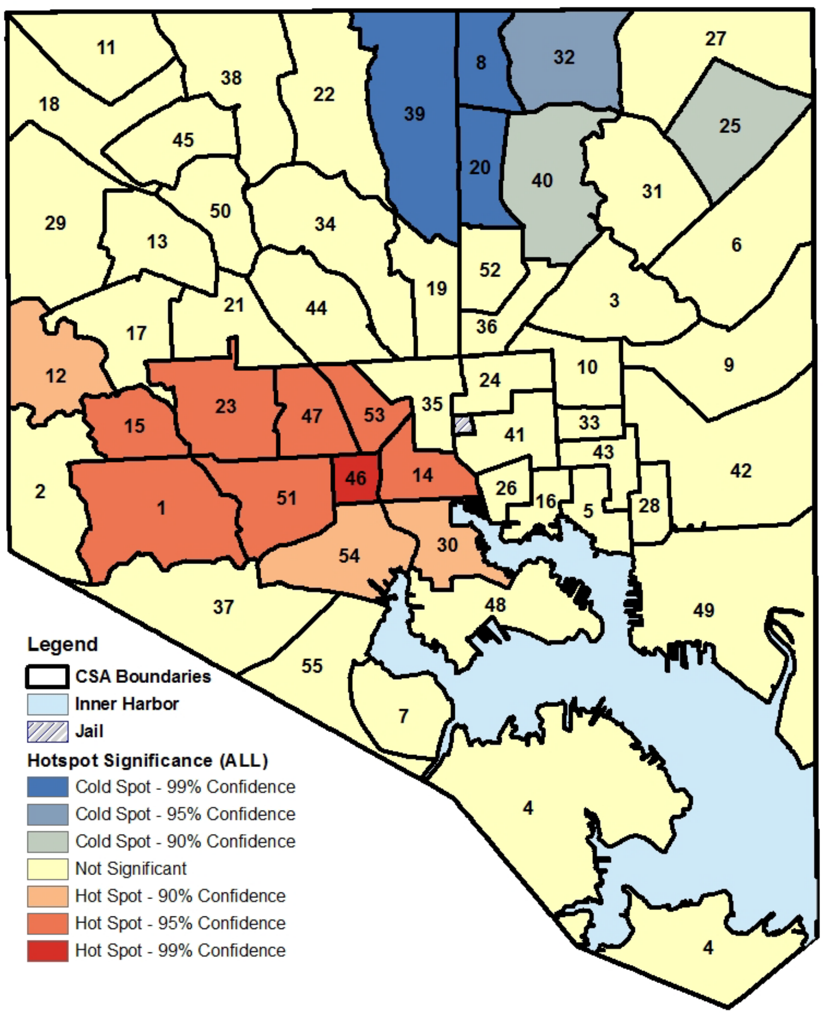


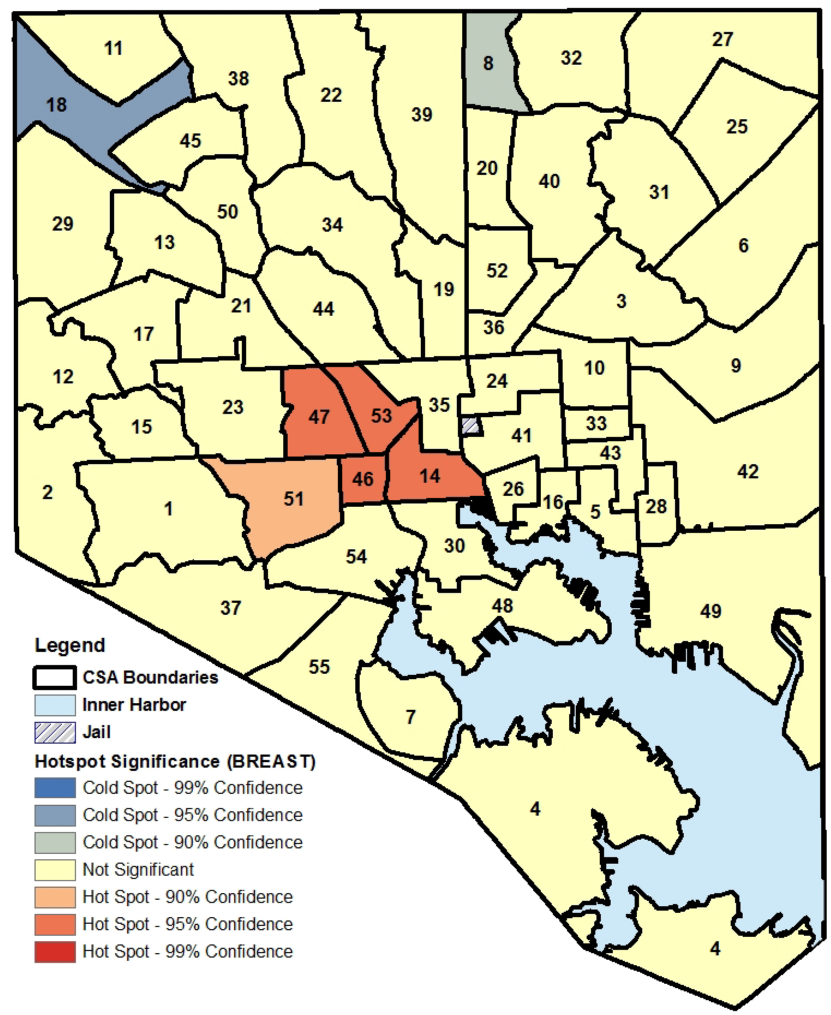


Appendix 2G Figure S2.8c-d: CSA distribution shaded by quintile of cancer stage (cervical and colorectal) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years

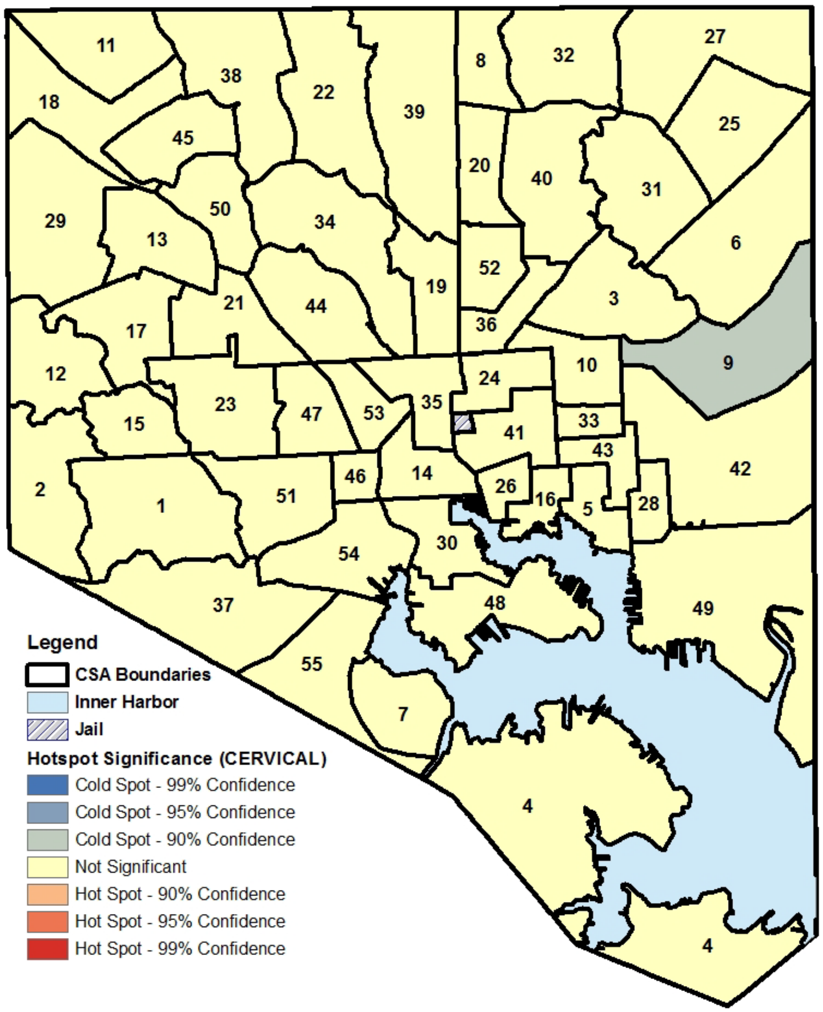


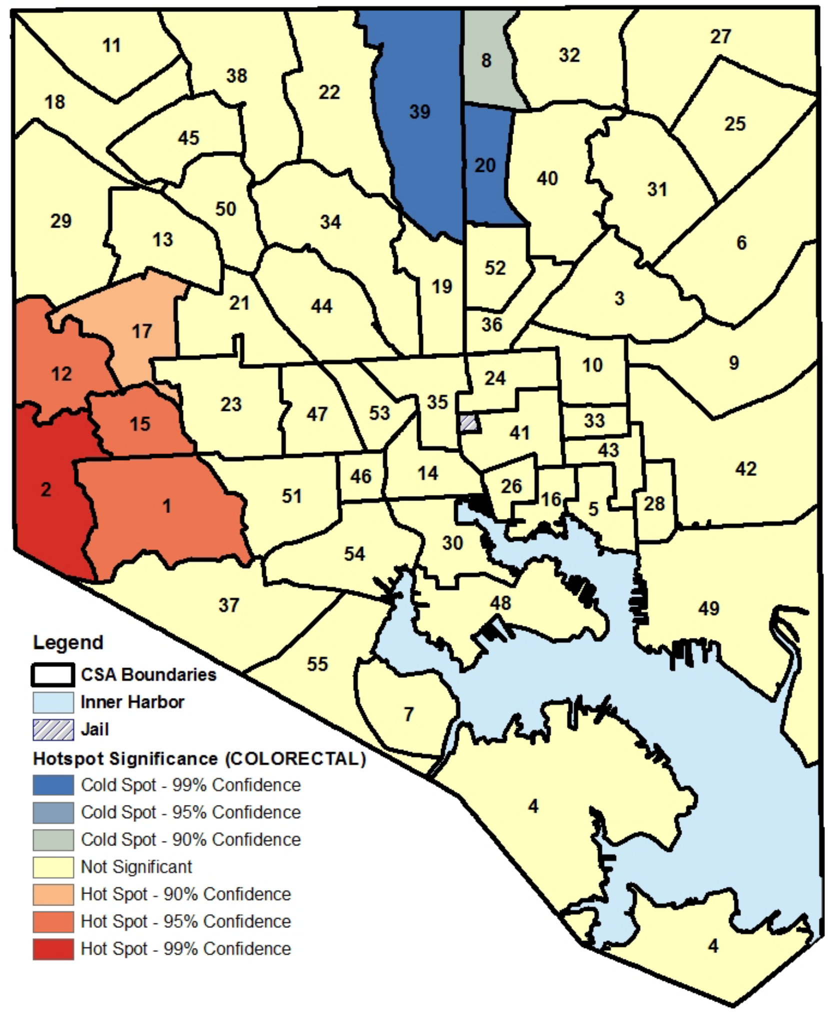
Appendix 2G Figure S2.9a-b: Hot spot analysis and statistical significance of cancer stage (all and breast) per 1,000 female residents aged 50 to 74 years by CSA





Appendix 2G Figure S2.9c-d: Hot spot analysis and statistical significance of cancer stage (cervical and colorectal) per 1,000 female residents aged 50 to 74 years by CSA





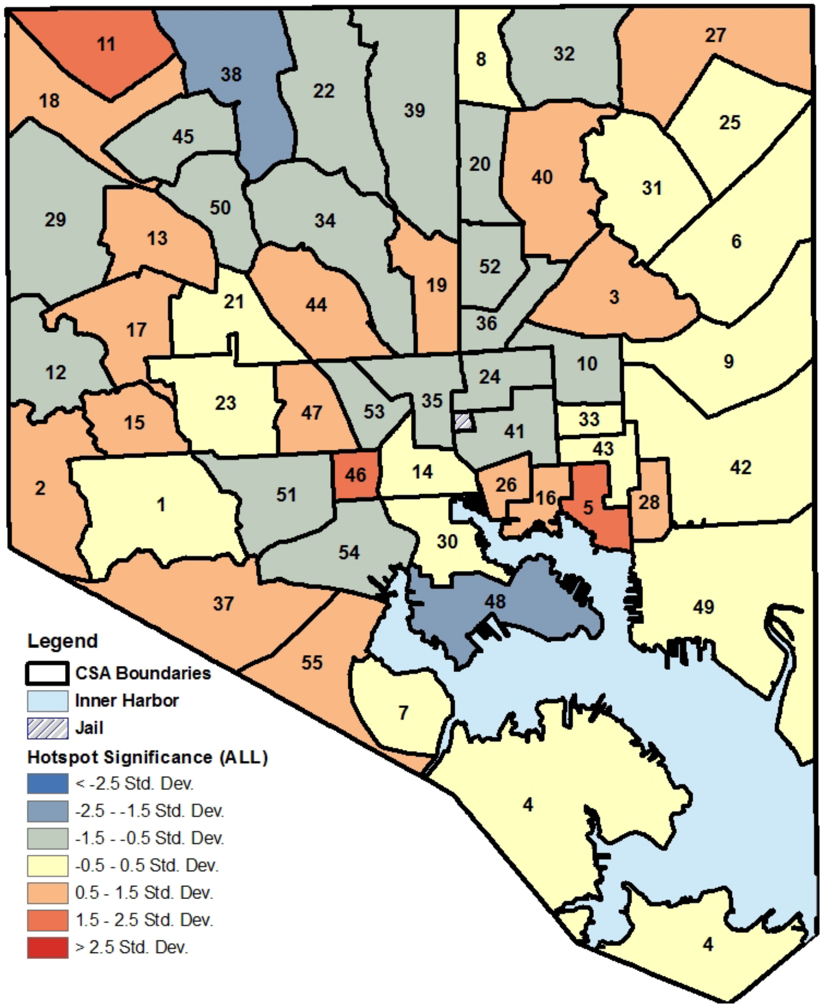
Appendix 2G Table S2.4: Ordinary Least Squares regression models for cancer stage among females 50-74 years by site and candidate neighborhood-level covariates

	All Cancers		Breast		Cervical		Colorectal	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Unadjusted								
Females								
50-74 yrs	0.000003	0.883	0.000003	0.888	0.000	0.997	-0.0001	0.131
% AA	0.001	0.012*	0.001	0.014*	-0.001	0.710	0.001	0.341
Racial Diversity	-0.0005	0.444	0.0004	0.506	0.001	0.693	-0.002	0.108
Household income <25K	0.003	0.006*	0.003	0.011*	-0.005	0.200	0.002	0.269
Female headed	0.002	0.003*	0.002	0.001*	-0.004	0.129	0.001	0.417
Vacants	0.003	0.026	0.002	0.108	-0.008	0.080	0.003	0.326
Housing violations	0.023	<0.001*	0.021	0.002*	-0.014	0.573	0.014	0.296
Crime	0.0002	0.389	0.0001	0.514	0.001	0.034*	0.0000	0.964
Domestic violence	0.002	0.005*	0.002	0.009*	-0.004	0.236	0.002	0.254
Teen births	0.001	<0.001*	0.001	0.0005*	-0.0001	0.961	0.001	0.276
Employed	-0.002	0.020	-0.003	0.019*	0.005	0.261	0.0002	0.943
Businesses	-0.00001	0.623	-0.00001	0.631	0.0002	0.012*	-0.0001	0.312
Voted	-0.003	0.018*	-0.004	0.013*	-0.0003	0.960	-0.001	0.689
Dirty streets	0.0004	0.051	0.0003	0.285	0.0004	0.051	0.0004	0.344
Tree coverage	-0.002	0.026*	-0.002	0.058	-0.001	0.658	0.001	0.450
Neighborhood associations	0.003	0.175	0.003	0.214	0.007	0.449	-0.003	0.595

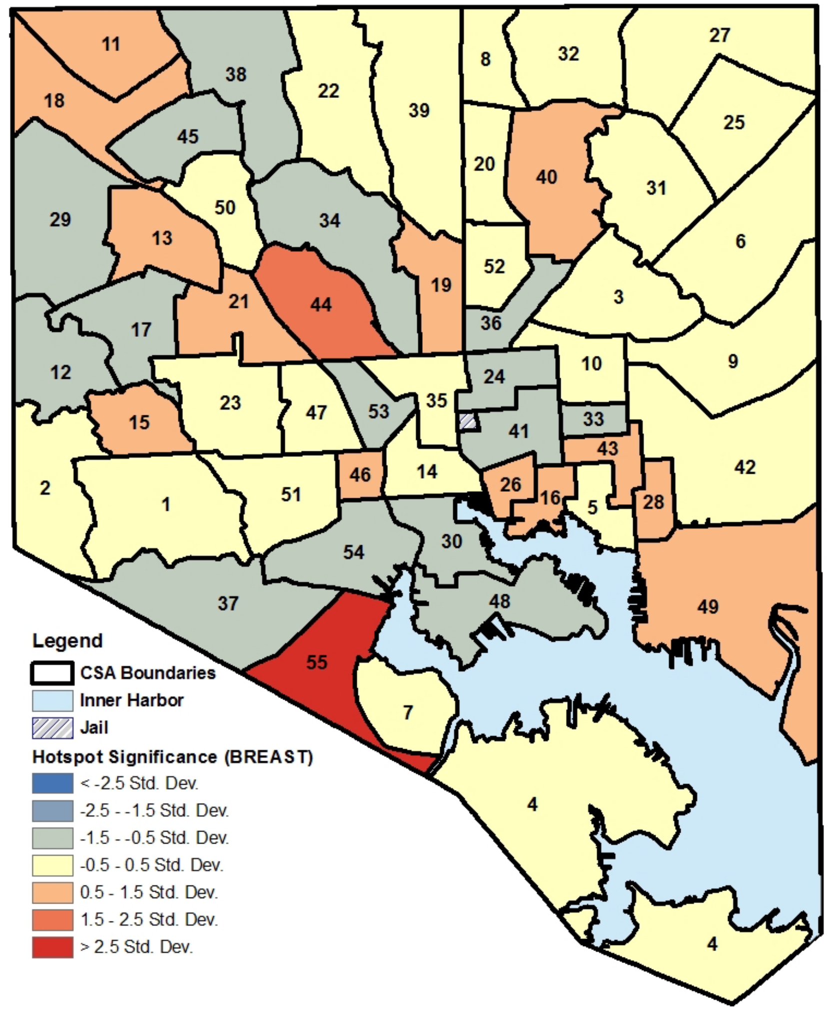
Adjusted	All Cancers Coefficient	p-value	Breast	Coefficient	p-value	Cervical	Coefficient	p-value
Females			Females			Females		
50-74 yrs	-0.00001	0.709	50-74 yrs	0.00001	0.731	50-74 years	0.00001	0.913
% AA	0.0001	0.828	% AA	-0.0002	0.84	Crime	-0.0005	0.798
Household income <25K	-0.001	0.576	Household income <25K	-0.0007	0.742	Businesses	0.003	0.226
Female headed	-0.0003	0.872	Female headed	0.003	0.203			
Housing violations	0.023	0.016*	Housing violations	0.014	0.231			
Domestic violence	-0.001	0.491	Domestic violence	-0.003	0.210			
Teen births	0.001	0.016*	Teen births	0.001	0.041*			
Voted	0.001	0.748	Employed	0.002	0.408			
Tree coverage	-0.0004	0.548	Voted	-0.001	0.685			
R-squared	0.287			0.174			0.097	

* Statistically significant

Appendix 2G Figure S2.10a-b: Spatial output of final models for ordinary least squares regression for cancer incidence (all and breast) of females aged 50 to 74

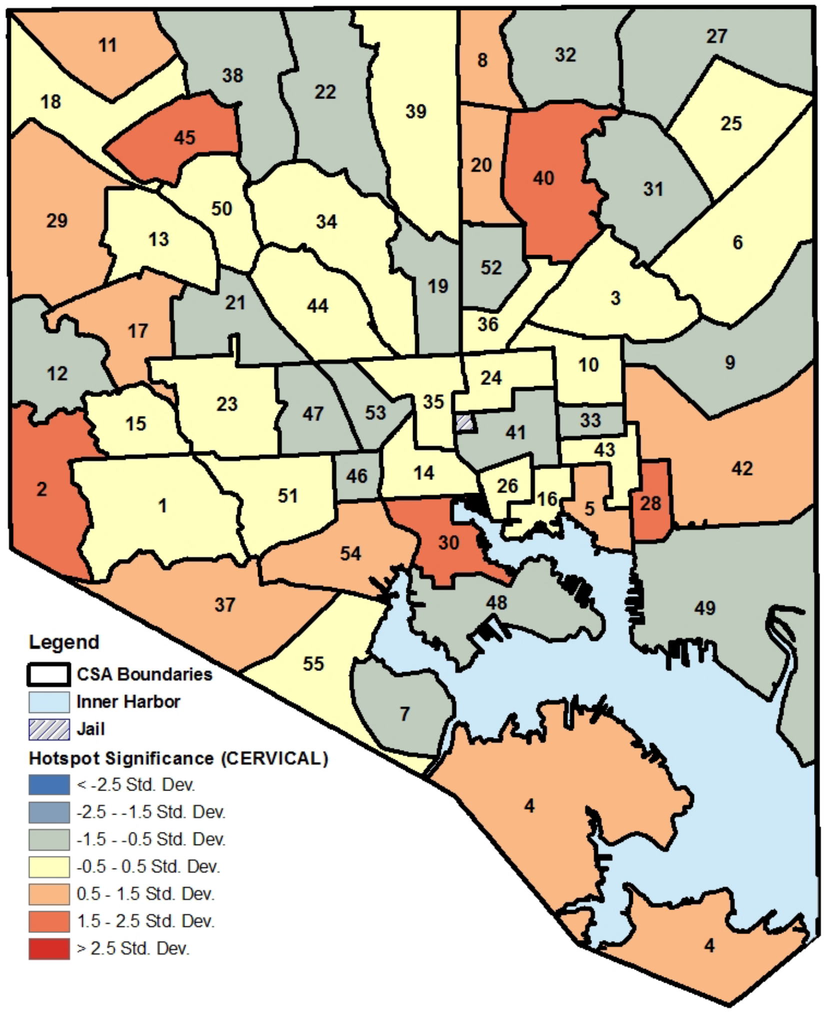


[a]Overall cancer model: Females (50-74 years), % African-American, household income, female-headed, housing violations, domestic violence teen births, voted, tree coverage



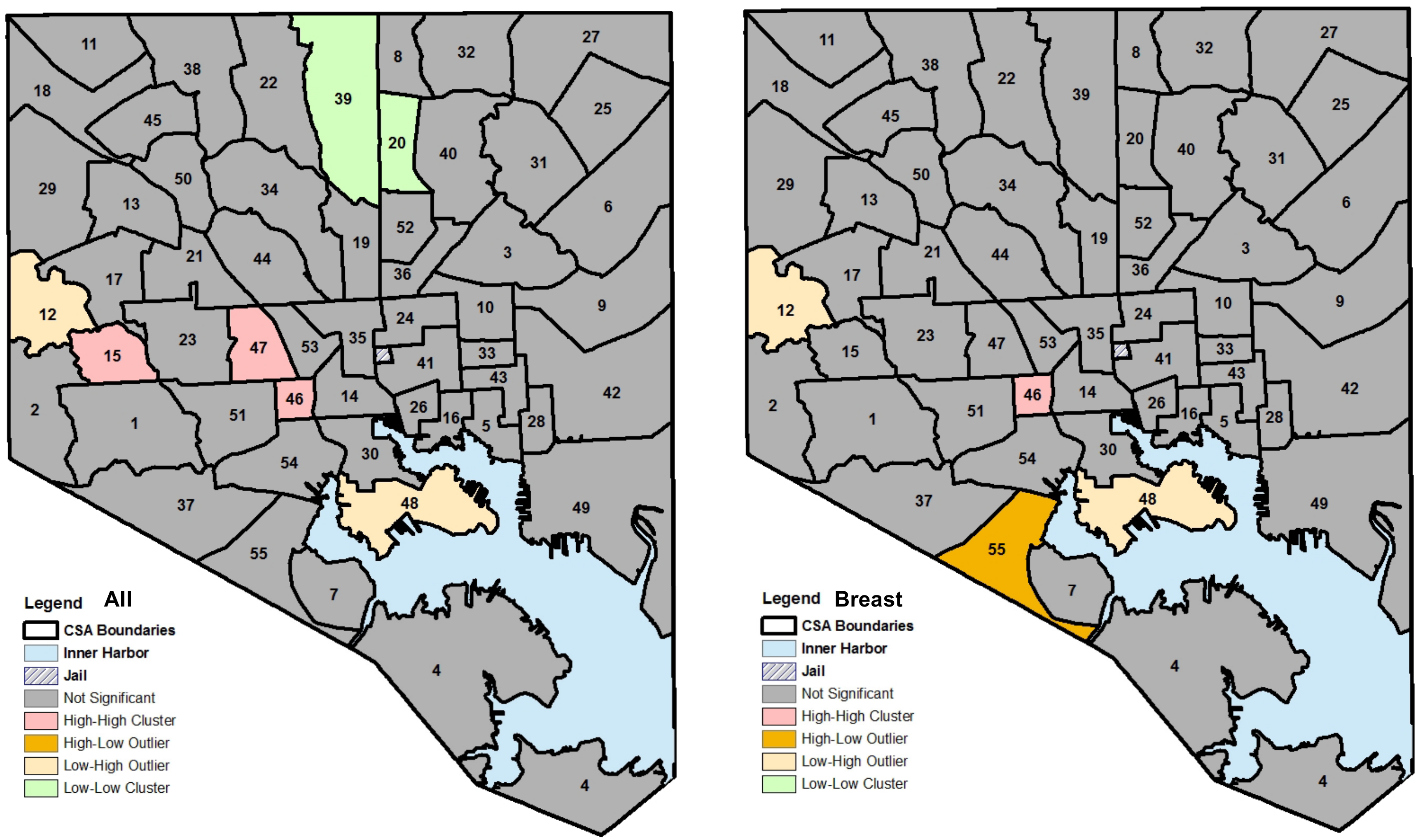
[b]Breast cancer model: Females (50-74 years), % African-American, housing violations, domestic violence, teen births, voted, tree coverage

Appendix 2G Figure S2.10c: Spatial output of final models for ordinary least squares regression for cancer incidence (cervical) in females aged 50 to 74 years

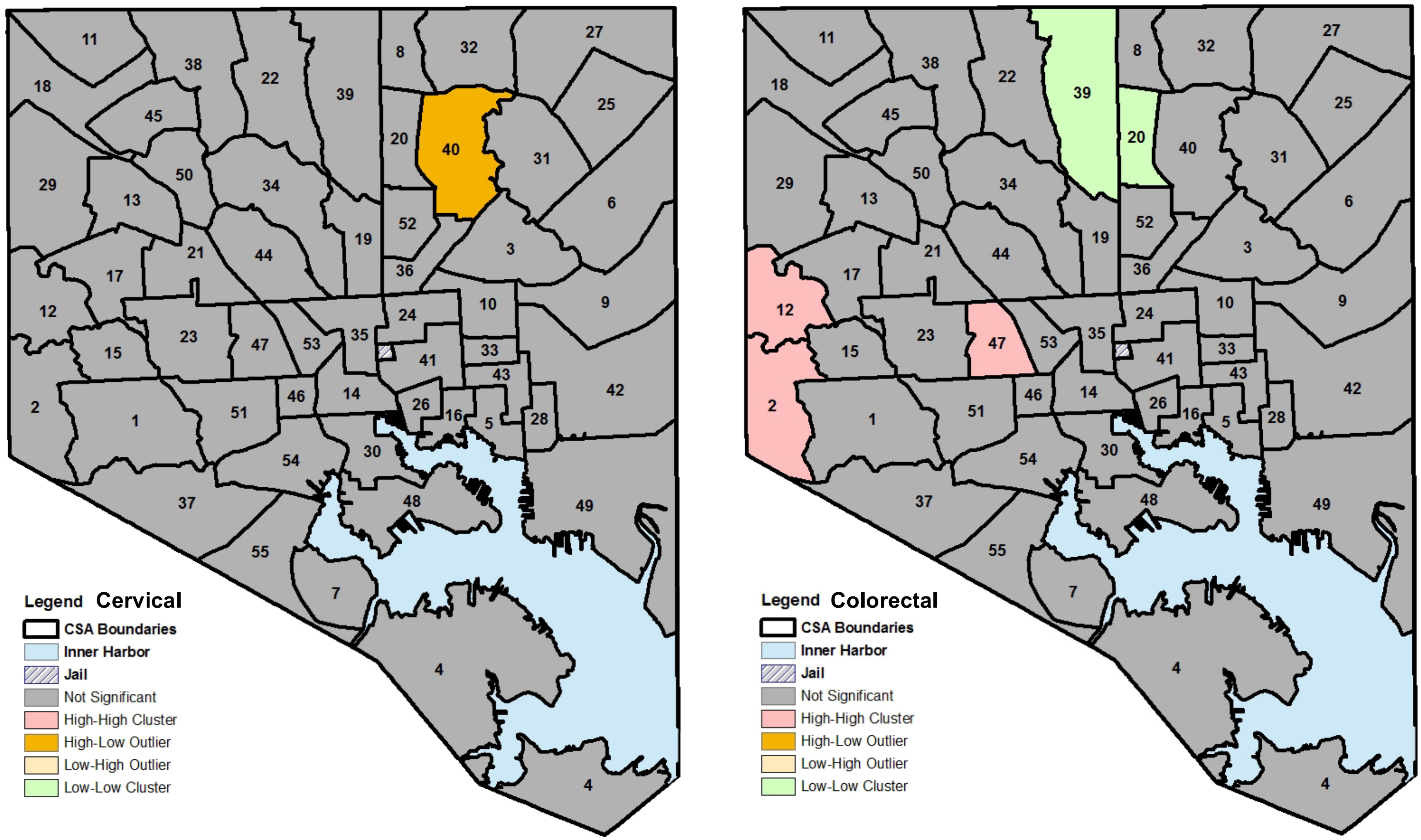


[c]Cervical cancer model: Females (50-74 years), crime, businesses

Appendix 2G Figure S2.11a-b: Local Moran's I analysis for cancer stage (all and breast) in females aged 50 to 74 years



Appendix 2G Figure S2.11c-d: Local Moran's I analysis for cancer stage (cervical and colorectal) in females aged 50 to 74 years



CHAPTER 3: EVALUATING NEIGHBORHOOD CORRELATES AND GEOGRAPHIC DISTRIBUTION OF BREAST, CERVICAL, AND COLORECTAL CANCER MORTALITY

Abstract
Objective: To evaluate the geographic variation of breast, cervical, and colorectal cancer mortality in Baltimore City and the neighborhood characteristics associated with the observed variation.

Design and Methods: Using female cancer cases that were identified by the Maryland Cancer Registry as being diagnosed with breast, cervical, or colorectal cancer in 2000 to 2010 while living Baltimore City, the geographic distribution of cancer deaths (N= 1,765) was evaluated for each cancer site. The cohort was restricted to females that had been diagnosed between the ages of 21 to 74 years. Geographical variation was ascertained through several cluster detection methods, which included choropleth maps, the Getis-Ord Gi* statistic, and the local Moran's I. Cancer mortality was defined for each cancer site as the number of deaths due to the primary site over the length of follow-up from the date of diagnosis to the date of either death or censoring.

Results: Breast, cervical, and colorectal cancer mortality each had noticeable geographic variation across Baltimore City. There were differences in terms of what areas had the highest absolute burden of mortality depending on the cancer site being evaluated. The three cancer sites rarely aggregated in the same communities. When attempting to explain these local distributions, the significantly associated neighborhood covariates also varied by cancer. Almost none of the neighborhood-level covariates explained the distribution of cervical and colorectal cancer mortality. However, breast cancer mortality was highly associated with several of the variables. Its final adjusted model explained 46.1% of the geographic variability observed.

Conclusions: It is important to evaluate cancer mortality at the local level to inform the efficient allocation of screening and treatment resources. There should also be discussions on improving data collection methods to appropriately capture cancer treatment at the community-level. There is currently not a scalable solution in place to efficiently document whether treatment patterns are affected by neighborhood context.

Introduction

While comprising nearly 30% of all cancer deaths in the United States, breast, cervical, and colorectal cancer have consistently yielded three of the top ten age-adjusted cancer death rates within the female population.[1,2] Although cervical cancer has maintained stable mortality from 2000 to 2009, diagnosis of disease at earlier and more treatable stages as well as therapy improvements have resulted in declining trends for breast and cancer deaths over that same time period.[3] Unfortunately, African-American females have experienced smaller declines in cancer mortality across cancer sites while even exhibiting an increase in cervical cancer death.

At the individual-level, the mechanisms behind this disproportionate burden of cancer mortality in racial-ethnic minorities have been widely studied. The stage at diagnosis has been known to be one of the strongest predictors of cancer prognosis. African-American women tend to get diagnosed with breast, cervical, and colorectal cancer at more invasive stages than their white counterparts.[3,4] Additionally, the administration of timely and appropriate treatment can greatly reduce the chance of death. Study results have consistently shown that African-Americans have less access to adjuvant treatment and are also more likely to attend under-resourced facilities.[5–9]

Recently, the literature on these individual-level predictors has been augmented by a recent shift in focus to the role of place on cancer mortality. This growing field has looked at the general geographic distribution of cancer mortality as well as the effect of the neighborhood's social conditions on the disease spectrum.[10–16] This movement towards evaluating place and its effect on health outcomes may uncover previously unseen epidemiological associations at the individual-level and also provide more upstream entry into the disease pathway. There has been increasing evidence that biological-environmental interactions can modify exposure-outcome relationships and lead to disparate health outcomes.[17]

This chapter aims to understand the geographic distribution of breast, cervical, and colorectal cancer mortality among female residents of Baltimore City, Maryland as well as the neighborhood characteristics associated with that distribution. The methodology to achieve this objective will integrate unique area-level measures that capture the social and physical conditions of Baltimore City neighborhoods. The overall purpose of this chapter is to provide an initial understanding as to the neighborhood context in which cancer mortality tends to occur more commonly in Baltimore City.

Methods

Cancer death ascertainment

The Maryland Cancer Registry (MCR), discussed in Chapter 1, was the data source for the primary outcome of cancer-specific mortality. The MCR obtained mortality data from other sources, such as the Maryland Vital Statistics Administration and National Death Index, through data sharing agreements. The cancer cohort from Chapter 2 (female Baltimore City residents diagnosed with breast, cervical, or colorectal cancer between the ages of 21 to 74 years from 2000 to 2010) was used as the source population for cancer deaths for this analysis. Individuals were excluded from the final sample if they had an unknown cause of death or if the date of death only had year-level granularity.

Neighborhood characteristics

The primary independent variables were the neighborhood-level characteristics made publicly available by the Vital Signs report for each Community Statistical Area (CSA) in Baltimore City.[18] The methodology of assessing the correlation among these covariates within and across domains as well as reducing the list to key covariates was described in Chapter 2. The CSA-level covariates were averaged across the years available based on the longitudinal consistency observed in the trend line plots created in Chapter 2.

Statistical analyses

Descriptive statistics were conducted across all deaths in the study sample of female residents with a breast, cervical, or colorectal diagnosis as well as stratified by site. For individual-level covariates provided by the MCR, mean age at death was calculated as well as percent distribution of tumor grade, race, cause of death, and treatment distributions within a cancer site. This set of initial statistics was carried out using the software STATA 12.1.[19]

At the CSA-level, the total number of cancer deaths within each geographic boundary was determined, and choropleth maps were created to shade CSAs by quintiles for cancer mortality. When ascertaining the event of interest, the primary outcome was defined as having one of the three targeted cancer sites and subsequently dying due to that same cancer. For example, a woman diagnosed with colorectal cancer was considered to have had an event, in this case death, if the cause of death was colorectal cancer. Individuals with a cause of death that was not the same as the primary site were censored on their date of death. For deaths with only month-level granularity available, these dates were assigned to mid-month. The outcome of mortality was calculated as a rate taking the number of cancer deaths due to the primary site within a CSA per person-years of follow-up, which was accrued from the date of diagnosis to the date of either death or censoring. Since deaths are added to the National Death Index approximately 12 months after the end of the calendar year in which the death took place, the dataset was assumed to have included all deaths that occurred up through July 31, 2013. Individuals that were identified as still being alive by the MCR were administratively censored on this date.

Similar to the methodology outlined in Chapter 2, the CSAs were shaded by quintiles in thematic maps to provide an exploratory visualization of what areas in Baltimore City had a higher absolute burden of cancer mortality relative to the percentile rank. The statistical significance of these potential spatial clusters was evaluated through the Getis-Ord Gi* statistic, which highlighted hot and cold spots. This spatial methodology determined whether the local cancer mortality, consisting of a CSA and its immediate neighbors, was more or less than expected as compared to the overall cancer mortality in Baltimore City.[20] This spatial assessment identified smaller geographic units of high burden that would have otherwise gone unnoticed if cancer mortality had only been reported at the county-level for Baltimore City.

The spatial cluster analysis was followed by the utilization of a global ordinary least squares (OLS) regression model for each cancer site to determine which of the CSA-level independent variables explained the distribution of the hot and cold spots for cancer mortality in Baltimore City. A p-value of <0.05 in these models would indicate that a neighborhood-level characteristic was associated and tracked with how the outcome varied geographically. The R-squared statistic for each of the cancer site-specific models was also noted, which provided the proportion of the geographic variation that was explained by the adjusted model.

As was done in Chapter 2, a local indicators of spatial association (LISA) function using local Moran's I was again conducted. This was carried out to determine if, and where, there were discordant clusters of cancer mortality. The LISA analysis identified how each CSA contributed towards the global magnitude of cancer mortality while also highlighting cluster outliers.[21,22] These outliers provide unique information by indicating CSAs that have, for example, high cancer mortality but have a neighboring CSA with low cancer mortality. Closer examination of the reasons behind this discordance might yield information that could be utilized for intervention development. The spatial methodology described above was conducted using ArcGIS 10.3 software.[23]

The Johns Hopkins School of Public Health's Institutional Review Board and the Maryland Department of Health and Mental Hygiene determined this study to be exempt research.

Results

Neighborhood characteristics

As previously described in the neighborhood results discussed in Chapter 2, multiple years of data made available by the Vital Signs report were condensed to a single summary measure for each indicator. This was based on the observation from trend line plots that the characteristics did not significantly change over time. The same reduced list of covariates described in Chapter 2 was utilized for this cancer mortality analysis.

Cancer mortality and population characteristics

From the original cohort used in Chapter 2, cases with unknown cause of death (n=25) or only a year for death date (n=12) were excluded. The final study population had 1,765 deaths across the three cancer sites: breast (n= 1,012), cervical (n= 198), and colorectal (n= 555). Person-time was accrued from date of diagnosis to either date of death due to the diagnosed cancer or censoring. The person-years accrued by cancer site were as follows: breast (22,356 years), cervical (1,844 years), and colorectal (5,734 years). Cervical cancer cases tended to be the youngest at time of death with an average age of 55.5 (SD=12.5) years compared to breast and colorectal cancer (**Table 3.1**). Overall, the majority of the cancer deaths were due to the primary cancer site.

The distribution of race among cancer deaths did not differ significantly when compared across cancer sites. However, the rest of the descriptive characteristics, such as stage and age at death, significantly differed across cancer sites. For example, colorectal cancer deaths had a distribution that skewed towards being a stage 3 or 4 at the time of diagnosis. Variations were also observed by treatment type; however, the MCR does not follow-up on initiated treatment after obtaining the reported diagnosis information. As a result, this data field had a high degree of missingness as cases with no treatment information ranged from 20% to 30%. Given the critical role that treatment plays in prognosis, this analysis was unable to evaluate whether treatment patterns or timeliness of care after a cancer diagnosis varied by neighborhood or cancer site. When evaluating total deaths by CSA, the CSA of Cedonia/Frankford (n=70) yielded the most cancer deaths while Poppleton/The Terraces/Hollins Market had the highest death rate at 75.6 cancer deaths due to the primary site per 1,000 person-years of follow-up (**Table 3.2**).

Spatial analysis of cancer mortality

The CSA map key was reproduced from Chapter 2 to assist in identifying the CSAs by name in the subsequent maps (**Table 3.3**). The initial distribution of cancer mortality, as shown in the choropleth maps, illustrated higher death rates in central and west Baltimore City (**Figures 3.1a-d**). Mortality appeared to aggregate in the same areas across the cancer sites. These preliminary visualizations provide early evidence of neighborhoods with a higher burden of mortality. Even without statistical significance, these initial exploratory mappings can prompt discussion as to the allocation of treatment resources.

While the analysis for this chapter used 21 to 74 years as the range for age at diagnosis, the methodology was also reproduced after restricting the population to an older age bracket of 50 to 74 years. The results for this additional geospatial analysis utilizing a more narrow age range can be found in **Appendix 3A** for comparison purposes and as an additional sensitivity analysis. The findings for the analysis consisting of the more narrow age range aligned very closely to the findings of the broader age interval of 21 to 74 years, which will be discussed in detail in this results section. Overall, the inferences and identified spatial clusters did not drastically change when using an older population. Although age is a strong risk factor for mortality, the 21 to 74 years range was used from this point forward for the primary analysis because screening guidelines, particularly for cervical cancer, apply to women within this age group.

The choropleth maps were also utilized to display cancer mortality by Z-scores across the neighborhoods. This mapping was carried out to improve the ability to make comparisons across each of the maps by cancer site. Upon applying the Z-score categorization to the color gradient, the number of CSAs exhibiting higher cancer mortality noticeably reduced across all three sites and the distribution of deaths became more similar. The thematic maps utilizing Z-scores across the CSAs for cancer mortality can be found in **Appendix 3B**.

The above assessment was followed up with the hot spot analysis, which evaluated the statistical significance of the exploratory aggregations observed in the quintile color gradient from the thematic maps. Similar to what was seen in the previous chapter where distribution varied by cancer, the hot spot analysis, which used mortality rates, noticeably differed by the primary diagnosis site (**Figures 3.2a-d**). Of the three cancer sites, only colorectal cancer yielded a substantial number of cold spots or places that had less than expected deaths. Breast cancer had a single cold spot, which directly overlapped with a cold spot for colorectal cancer mortality.

Using the same approach as Chapter 2, a heat map was provided to more easily identify whether CSAs consistently ranked in the 80th percentile for cancer mortality across sites of diagnosis (**Table** 3.**4**). From this visualization it can be observed that only 2 of the 55 CSAs had mortality fall within the same quintile across each of the primary tumors. More importantly, this rendering allows for the quick identification of CSAs that yielded the highest mortality quintile for one cancer type but the lowest mortality quintile for another. This was the case for CSAs such as Cross-Country/Cheswolde, Greater Govans, and Orangeville/E. Highlandtown.

In order to explain the drivers behind these hot and cold spots, an unadjusted OLS regression model was conducted for each of the CSA-level characteristics from the reduced list of covariates. The unadjusted associations with cancer mortality can be found in **Table 3.5a**. Many of the covariates were significantly associated with breast cancer mortality. For the other two cancer sites of interest, only teen births demonstrated a significant association with colorectal cancer mortality. The adjusted models included all covariates that were statistically significant at the univariate level as well as age due to its known relationship with mortality. While the majority of the neighborhood-level characteristics did not retain their statistical significance within the adjusted model, the final model for breast cancer explained 46% of the geographic distribution observed for mortality (**Table 3.5b**). The lack of more significant associations for colorectal cancer could be explained by two possibilities. First, it could be that there is truly no association between community-level covariates and colorectal mortality. However, the more likely scenario is that none of the available indicators from the Vital Signs report appropriately captures a social infrastructure mechanism that would affect colorectal cancer mortality. Meanwhile, the absence of significant associations for cervical cancer might be attributed to its low sample size.

When compared to the final models of Chapter 2 (cancer incidence), the neighborhood-level covariates explained more of the variation in cancer mortality than in cancer incidence, specifically when looking at breast cancer mortality, which had nearly half of its variability, as shown by the R-squared statistic (R^2= 46.1%), explained. This explained variation ties into the results seen in the mapping of the residuals for the final models (**Figures 3.3a-c**). For the most part, the residuals appear to be distributed randomly across the map with no major clustering. Residuals are indicators of how well the model fits or predicts the association being evaluated. A model that fits well will have smaller residuals since the predicted values align closely with the observed values.

It is important that residuals are normally distributed to provide evidence that the model is unbiased. An aggregation of geographic residuals in certain areas of Baltimore City would be an indicator that the model is differentially fitting certain CSAs better than others. There was some aggregation of cold spots for colorectal cancer in the spatial output of its adjusted model. As already discussed, this could be due to the lack of appropriate CSA measures that would better explain the geographic distribution of colorectal cancer mortality.

The local Moran's I results presented some low-high and high-low clusters for each of the cancers in different areas (**Figure 3.4a-d**). Both breast and colorectal cancer mortality had low-high clusters in the same vicinity around the harbor area. Cervical also yielded one high-low cluster, which incidentally overlapped with a high area of colorectal deaths. These discordant clusters require more in-depth research to explain what immediate feature is protecting or harming these CSAs relative to their neighbors.

Conclusion

The three main results provided in this chapter were: 1) the varying burden of cancer mortality across Baltimore City neighborhoods and cancer sites; 2) the differences in the neighborhoods that yielded the highest rates of cancer mortality for breast, cervical, and colorectal cancer; and 3) the differences in the observed associations between the cancer site-specific mortality and the neighborhood characteristics explaining the observed geographic distribution. As hypothesized before the study, the choropleth visualizations and identification of geospatial clusters confirmed that cancer mortality is not randomly distributed across Baltimore City. There are clear area-level pockets that experience a higher burden of cancer death.

While there is limited literature on the topic, these findings are consistent with what has been produced thus far in terms of geographic variation in cancer mortality. Prior studies have noted that mortality is not randomly distributed with a geographic region and that oftentimes the observed non-random distribution aligned with area-level measures, such as socioeconomic status and neighborhood deprivation.[11-13] The improvements this dissertation makes on those analyses are that it integrates measures that capture the social and physical conditions of neighborhoods in a more in-depth way. This comprehensive set of indicators allows for the opportunity to capture relationships that may have otherwise been overlooked using more traditional Census data. For example, the indicator describing the proportion of low-income residents did not yield a significant association with the geographic variation observed for cancer mortality in this analysis. However, significant relationships were observed with the unique area-level characteristics available through the Vital Signs report. As a result, the findings are uncovering relationships that had previously gone unevaluated due to the lack of these community-level data points.

This overall finding demonstrates the necessity to create interventions that are geographically tailored in order to most efficiently utilize limited resources for screening and treatment. By not keeping place in mind, cancer programs or treatments run the risk of being deemed ineffective if they are directed towards areas with a lower burden of disease rather than to less healthy locations that might have yielded significant results. There is also a realistic possibility of actually contributing towards widening the gap in health inequities by directing resources towards areas of low disease.

It is worth noting there were differences in both the location of geospatial clusters for cancer mortality across cancer sites as well as in terms of which CSA-level characteristics yielded significant associations with the outcome. These differences are particularly relevant if a specific neighborhood covariate is being targeted upstream through an intervention. Due to these varying associations, improvements across all three cancer sites may not be observed if the community-level characteristics being targeted were only significantly associated with one of the primary cancers. These nuanced dissimilarities would be important to take into account when setting measures to ascertain effectiveness of programs.

This chapter contains a number of strengths and limitations, many of which overlap with those described in the previous chapter. Similar to Chapter 2, the concept of neighborhood is restricted to only where these individuals resided at the time of their diagnosis. Individuals are not only exposed to the physical and social conditions of their homes as they may socialize or be employed in other communities.[24] These other environments may positively or negatively affect their resources or social context perpetuated by their residential CSA.

On a related note, the restriction of only being able to extract an individual's address at the time of cancer diagnosis also meant that there was no information available for events that occurred between the time of their diagnosis and subsequent date of death. Cancer cases may have moved to either outside of Baltimore City or to another CSA. This relocation could have been caused by the diagnosis, such as changing residences to receive care elsewhere or move in with a caregiver. There has been little research published on residential mobility between cancer diagnosis and death. Most of the literature has focused on the role of mobility on cancer within the scope of the association between radon exposure and lung cancer.[25, 26] With respect to mobility in the general population of Baltimore City, the U.S. Census Bureau reported that over 80% of the study's target population of females aged of 21 to 74 years reported living in the same household as the previous year.[27] This percentage was consistent across the study period of 2000 to 2010 indicating a stability in residential status.

Another potential limitation is the possible misclassification of the outcome due to the utilization of cancer site-specific mortality. The number of deaths for each cancer site could have been miscounted if the cause of death was misclassified. The population had the majority of its deaths, ranging from 65% to 70% across the three cancers, caused by the primary site of interest. Evidence has shown that the accuracy of the reported cause of death is high among patients with distant stage disease and those with only a single type of primary cancer.[28] Nearly all of the cases in the study sample had a single primary cancer reported to the MCR. However, there was some variation across the cancer sites with respect to stage at diagnosis. As a result, accuracy of outcome classification may vary across cancer sites since the literature has indicated that later stage cancer is more likely to have death classified correctly as compared to early stage cancer.

Additionally, the MCR dataset did not have up-to-date treatment information on cancer cases. As a result, the analysis could not ascertain the geographic variation of treatment patterns across Baltimore City CSAs. In turn, it could not be assessed whether or not certain neighborhoods were more likely to have residents that had long lags between their diagnosis date and treatment initiation. As interest in the role of place on disease continues to grow, it might be worth re-evaluating the current data collection mechanisms in place at the MCR and whether treatment data points should be pursued with more rigor. This additional information could provide a better understanding of how quality of care can vary at the community-level.

Finally, the methodology presented in this chapter did not include individual-level characteristics and risk factors in its regression models. Their exclusion signifies an inability of the results to distinguish the degree of causality stemming from the CSA-level characteristics on cancer mortality versus individual-level covariates. While a more etiological-oriented research question would have required the addition of these characteristics, the intent of this particular study was to evaluate the burden of cancer death across Baltimore City neighborhoods. It is the evaluation of disease burden that meets the immediate needs of policy makers by steering where resources and services should be allocated. The main catalyst of this research was to answer the question of where cancer mortality tends to occur in order to propose informed and evidence-based solutions.

Overall, this chapter presented a population-based way to understand and identify the role of neighborhood context within the realm of cancer mortality disparities by taking the broader perspective of neighborhood context. Given the community-level exposures utilized in this study, cross-level inferences should not be made when interpreting the associations observed in the regression models. For example, CSAs with a high proportion of households earning an income of less than $25,000 were shown to have higher breast cancer mortality. This same inference should not be translated to the individual-level, which would result ecologic fallacy. A misinterpretation of this community association would be to assume the residents of these lower income households were the same ones dying due to breast cancer.

This research offers a number of strengths. As previously mentioned in Chapter 2, Baltimore City offers unique community-level characteristics that better capture the overall residential conditions of cancer cases. Most of the literature has relied heavily on U.S. Census data, which lacks the depth and breadth of the indicators provided by the Vital Signs report. Similarly, the setting of Baltimore City afforded the opportunity of using CSAs as the primary unit of analysis in the spatial methods of the study. This is relevant in two primary ways. First, it improved the granularity at which associations could be evaluated. Often, geographic variations might be overlooked as a result of using only large geographic units such as at the county or state-level. Secondly, it provides a meaningful neighborhood boundary for Baltimore City residents. Through the dissemination efforts of the Baltimore Neighborhood Indicators Alliance, residents have been exposed to these community designations and their corresponding indicators before. There is some high-level understanding among the residents as to the makeup of these CSAs.

Finally, this methodology integrated a transdisciplinary approach. There is usually a heavy focus on the traditional medical model and evaluation of etiological associations. This study's interest in community conditions and the impact of place combines several disciplines and areas of expertise. For the outcome of cancer mortality, the findings support the pursuit of more in-depth research into the communities with a high burden of death. It also supports the leveraging of these results into actionable agenda items for policy makers and cancer researchers.

References:

1. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2012 Incidence and Mortality Web-based Report . Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2015. Available from: www.cdc.gov/uscs

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015;65(1):5–29.

3. Jemal A, Simard EP, Dorell C, et al. Annual Report to the Nation on the Status of Cancer, 1975–2009, Featuring the Burden and Trends in Human Papillomavirus (HPV)–Associated Cancers and HPV Vaccination Coverage Levels. JNCI J Natl Cancer Inst 2013;:djs491.

4. Morris AM, Rhoads KF, Stain SC, Birkmeyer JD. Understanding racial disparities in cancer treatment and outcomes. J Am Coll Surg 2010;211(1):105–13.

5. Bickell NA, Shastri K, Fei K, et al. A tracking and feedback registry to reduce racial disparities in breast cancer care. J Natl Cancer Inst 2008;100(23):1717–23.

6. Ayanian JZ, Zaslavsky AM, Fuchs CS, et al. Use of adjuvant chemotherapy and radiation therapy for colorectal cancer in a population-based cohort. J Clin Oncol 2003;21(7):1293–300.

7. Baldwin L-M, Dobie SA, Billingsley K, et al. Explaining Black – White Differences in Receipt of Recommended Colon Cancer Treatment. J Natl Cancer Inst 2005 [cited 2015 Dec 10];97(16). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3138542/

8. Birkmeyer JD, Siewers AE, Finlayson EVA, et al. Hospital Volume and Surgical Mortality in the United States. New England Journal of Medicine 2002;346(15):1128–37.

9. Onega T, Duell EJ, Shi X, Demidenko E, Goodman D. Influence of place of residence in access to specialized cancer care for African Americans. J Rural Health 2010;26(1):12–9.

10. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. Am J Epidemiol 2002;156(5):471–82.

11. Singh GK, Miller BA, Hankey BF, Edwards BK. Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975–1999. Bethesda, MD: National Cancer Institute; 2003.

12. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. Br J Cancer 2002;87(6):585–90.

13. Klassen AC, Pankiewicz A, Hsieh S, Ward A, Curriero FC. The association of area-level social class and tobacco use with adverse breast cancer characteristics among white and black women: evidence from Maryland, 1992-2003. Int J Health Geogr 2015;14:13.

14. Gerend MA, Pai M. Social Determinants of Black-White Disparities in Breast Cancer Mortality: A Review. Cancer Epidemiol Biomarkers Prev 2008;17(11):2913–23.

15. Hiatt RA, Breen N. The Social Determinants of Cancer. American Journal of Preventive Medicine 2008;35(2):S141–S150.

16. Williams DR, Jackson PB. Social Sources Of Racial Disparities In Health. Health Aff 2005;24(2):325–34.

17. Warnecke RB, Oh A, Breen N, et al. Approaching Health Disparities From a Population Perspective: The National Institutes of Health Centers for Population Health and Health Disparities. Am J Public Health 2008;98(9):1608–15.

18. Baltimore Neighborhood Indicators Alliance–Jacob France Institute. Vital Signs 10: Measuring Baltimore's Continued Progress Towards Strong Neighborhoods and a Thriving City. Baltimore, MD: University of Baltimore; 2012.

19. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.

20. Ord JK, Getis A. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 1995;27(4):286–306.

21. Anselin L. Local Indicators of Spatial Association—LISA. Geographical Analysis 1995;27(2):93–115.

22. Moraga P, Montes F. Detection of spatial disease clusters with LISA functions. Stat Med 2011;30(10):1057–71.

23. ESRI. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute; 2011.

24. Riley RW. Health Starts Where We Learn . Robert Wood Johnson Foundation. [cited 2015 Nov 11]; Available from: http://www.rwjf.org/en/library/research/2010/10/health-starts-where-we-learn.html

25. Warner KE, Mendez D, Courant PN. Toward a more realistic appraisal of the lung cancer risk from radon: the effects of residential mobility. Am J Public Health. 1996 Sep; 86(9): 1222-7

26. Krewski D, et al. Residential radon and risk of lung cancer: a combined analysis of 7 North American case-control studies. Epidemiology. 2005 Mar; 16(2):137-45.

27. U.S. Census Bureau. Geographic Mobility in the Past Year by Age for Current Residence in the Unite States: 2006-2010 American Community Survey 5-Year Estimates. U.S. Census Bureau; 2010.

28. Lund JL, et al. Should Cause of Death From the Death Certificate Be Used to Examine Cancer-Specific Survival? A Study of Patients With Distant Stage Disease. Cancer Invest. 2010 Aug; 28(7): 758–764.

Tables and Figures:

Table 3.1: Descriptive characteristics of female cancer deaths in Baltimore City, MD

	Breast (n= 1,012)	Cervical (n= 198)	Colorectal (n= 555)	All sites (N= 1,765)	p-value
Mean age at death- years (SD)	60.6 (12.1)	55.5 (12.5)	64.2 (10.3)	61.1 (11.9)	<0.001*
Cancer stage- % (n)					
Stage 0	<0.6 (<6)	<3.0 (<6)	1.1 (6)	0.6 (11)	<0.001*
Stage I	29.2 (295)	19.7 (39)	16.4 (91)	24.1 (425)	
Stage II	14.4 (146)	15.7 (31)	10.1 (56)	13.2 (233)	
Stage III	28.6 (289)	31.3 (62)	20.4 (113)	26.3 (464)	
Stage IV	17.1 (173)	18.7 (37)	38.6 (214)	24.0 (424)	
Tumor grade- % (n)					
Grade I	7.7 (78)	6.1 (12)	6.5 (36)	7.1 (126)	<0.001*
Grade II	24.1 (244)	31.8 (63)	54.1 (300)	34.4 (607)	
Grade III	48.4 (490)	27.8 (55)	19.5 (108)	37.0 (653)	
Grade IV	1.1 (11)	<3.0 (<6)	<1.1 (<6)	1.1 (20)	
Cause of death- % (n)					
Primary site	65.5 (663)	64.1 (127)	68.8 (382)	66.4 (1,172)	<0.001*
Non-primary site	8.6 (87)	15.7 (31)	13.2 (73)	10.8 (191)	
Non-cancer	25.9 (262)	20.2 (40)	18.0 (100)	22.8 (402)	
Mean survival time[a] - years (SD)	2.8 (2.3)	1.6 (1.9)	2.0 (2.0)	2.4 (2.2)	<0.001*
Race					
White Non-Hispanic	24.7 (250)	24.8 (49)	27.9 (155)	25.7 (454)	0.234
Black Non-Hispanic	72.0 (729)	71.2 (142)	67.8 (376)	70.7 (1,247)	
Other	1.5 (15)	0.0 (0)	1.1 (6)	1.2 (21)	
Chemotherapy- % (n)					
Yes	43.7 (442)	49.5 (98)	39.5 (219)	43.0 (759)	<0.001*
No	30.8 (312)	28.8 (56)	32.43 (180)	31.1 (548)	
Hormone therapy- % (n)					
Yes	16.3 (165)	<3.0 (<6)	0.0 (0)	-[b]	<0.001*
No	57.7 (584)	77.3 (153)	79.3 (440)	66.7 (1,177)	
Radiation therapy- % (n)					
Yes	26.1 (264)	61.6 (122)	8.5 (47)	24.5 (433)	<0.001*
No	67.7 (685)	29.3 (58)	84.9 (471)	68.8 (1,214)	
Immunotherapy- % (n)					
Yes	1.0 (10)	0.0 (0)	1.3 (7)	1.0 (17)	0.323
No	82.6 (836)	81.8 (162)	79.8 (443)	81.6 (1,441)	
Surgery- % (n)					
Yes	70.2 (710)	33.3 (66)	71.2 (398)	66.5 (1,174)	<0.001*
No	22.7 (230)	56.6 (112)	20.2 (112)	25.7 (454)	

[a] From date of diagnosis to date of death due to primary site
[b] Data not presented to avoid back calculation
[c] Cells with fewer than 6 cases have been suppressed and indicated as “<6”
[d] Percentages may not add up to 100% due to missing data

Table 3.2: Community Statistical Area distribution of female cancer deaths in Baltimore City, MD

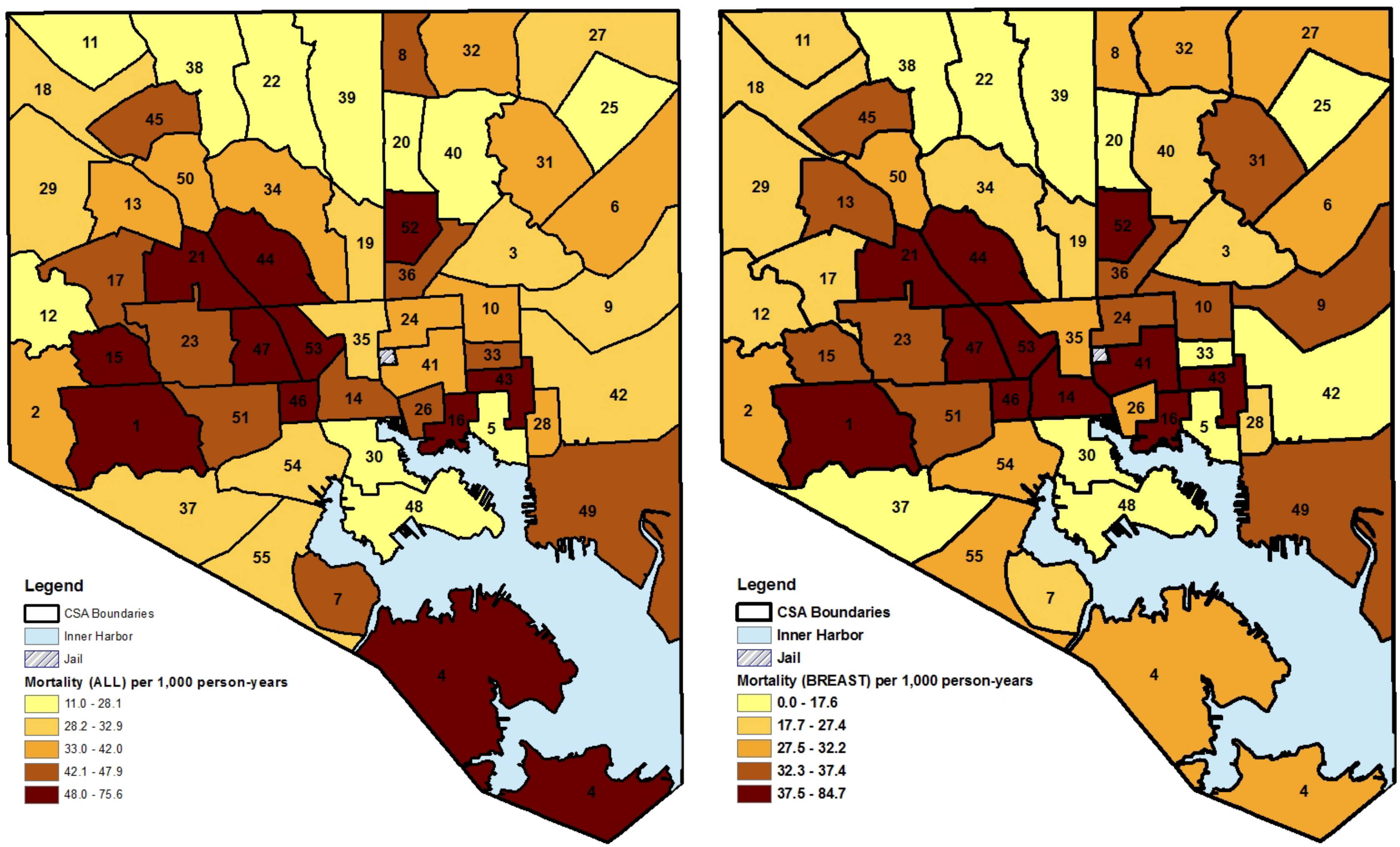
	Cancer deaths (n= 1,765)	Death rate[b] (per 1,000 PY)
Allendale/Irvington/S. Hilton	53	56.4
Beechfield/Ten Hills/West	32	42.0
Belair-Edison	38	32.9
Brooklyn/Curtis Bay	45	50.9
Canton	15	23.1
Cedonia/Frankford	70	35.5
Cherry Hill	20	45.0
Chinquapin Park/Belvedere	28	46.1
Claremont/Armistead	25	31.8
Clifton-Berea	31	39.7
Cross-Country/Cheswolde	24	25.0
Dickeyville/Franklintown	8	23.8
Dorchester/Ashburton	56	41.9
Downtown/Seton Hill	9	46.9
Edmonson Village	37	50.8
Fells Point	20	60.0
Forest Park/Walbrook	38	47.5
Glen-Fallstaff	44	31.0
Greater Charles Village/Barclay	29	32.2
Greater Govans	31	26.0
Greater Mondawmin	34	49.9
Greater Roland Park/Poplar Hill	8	11.0
Greater Rosemont	69	46.6
Greenmount East	39	38.2
Hamilton	33	24.3
Harford/Echodale	13	47.3
Highlandtown	31	28.9
Howard Park/West Arlington	13	38.8
Inner Harbor/Federal Hill	42	32.0
Jonestown/Oldtown	21	19.6
Lauraville	33	37.0
Loch Raven	49	34.9
Madison/East End	20	47.1
Medfield/Hampden/Woodberry/Remington	37	35.1
Midtown	30	30.4
Midway/Coldstream	39	46.6
Morrell Parkk/Violetville	27	30.6
Mount Washington/Coldspring	9	13.3
North Baltimore/Guilford/Homeland	30	11.4
Northwood	44	28.1
Orangeville/E. Highlandtown	40	39.8
Patterson Park North & East	24	29.1
Penn North/Reservoir Hill	50	60.3
Perkins/Middle East	21	64.8
Pimlico/Arlington/Hilltop	42	47.6
Poppleton/The Terraces/Hollins Market	15	75.6
Sandtown-Winchester/Harlem Park	66	63.6
South Baltimore	<6	13.2
Southeastern	23	47.9
Southern Park Heights	52	41.6
Southwest Baltimore	60	42.4
The Waverlies	25	53.6
Upton/Druid Heights	36	59.6
Washington Village/Pigtown	13	31.4
Westport/Mount Winans/Lakeland	19	31.1

[a] Cells with fewer than 6 cases have been suppressed and indicated as “<6”
[b] Defined as total deaths due to primary site per 1,000 person-years

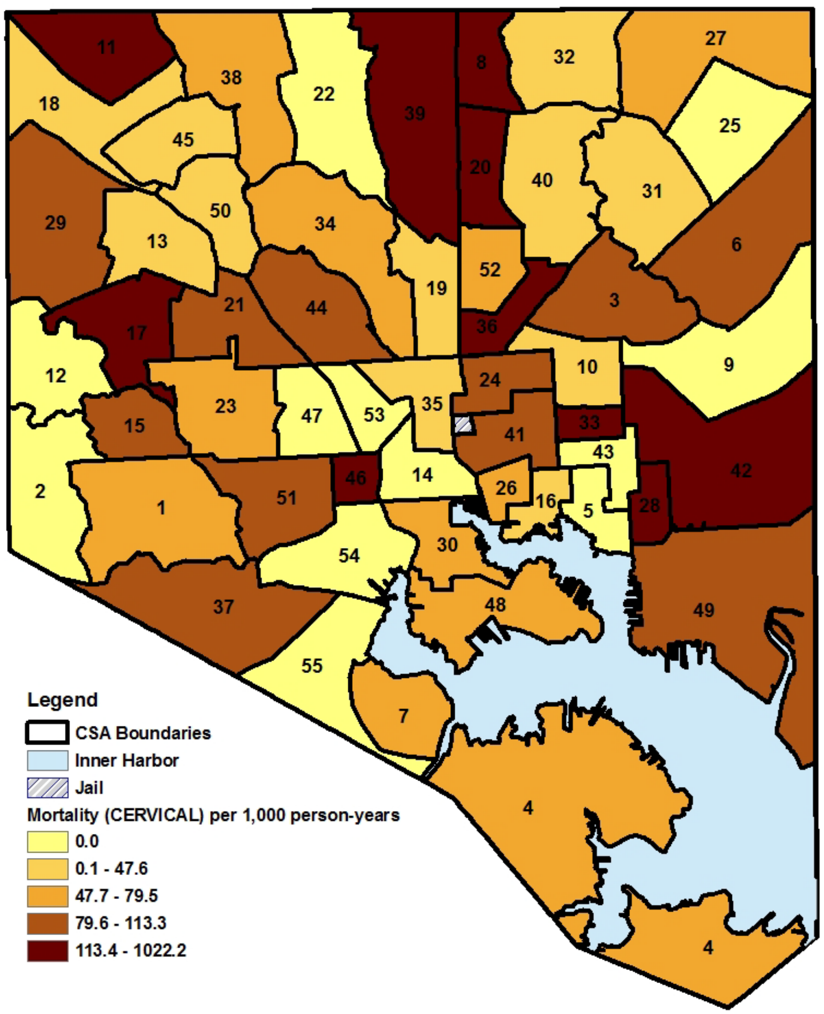
Table 2.4: Key for Community Statistical Area (CSA) map

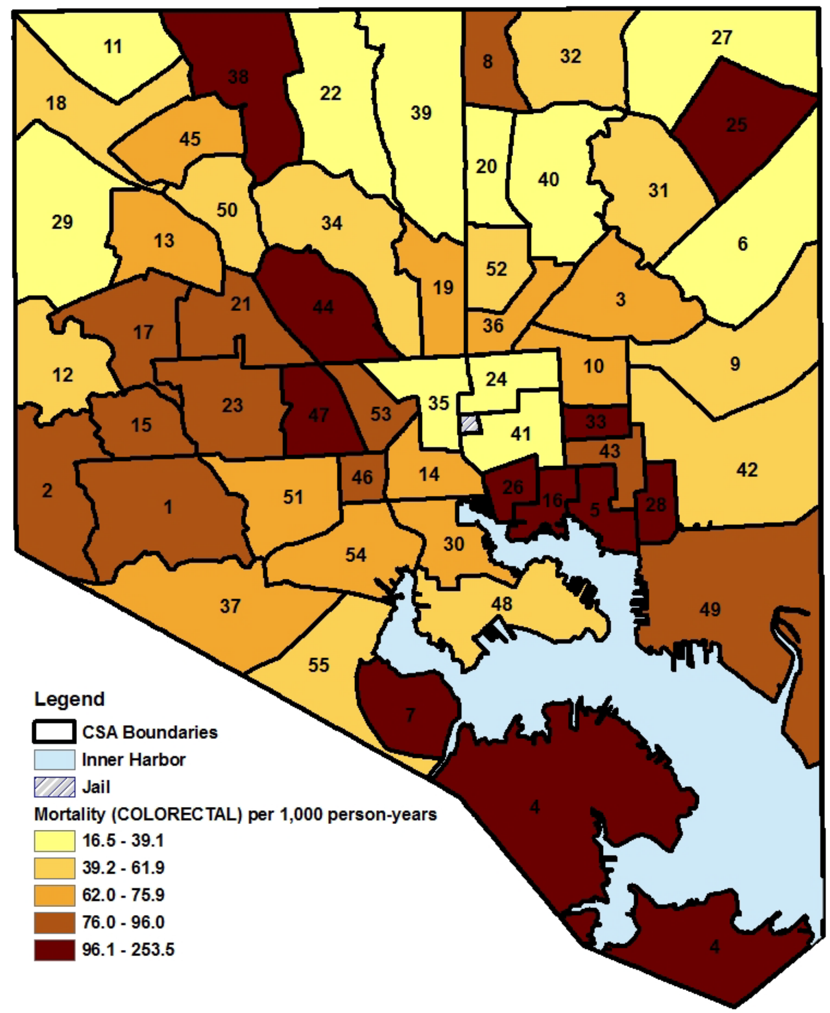
Allendale/Irvington/S. Hilton	1	Howard Park/West Arlington	29
Beechfield/Ten Hills/West	2	Inner Harbor/Federal Hill	30
Belair-Edison	3	Lauraville	31
Brooklyn/Curtis Bay	4	Loch Raven	32
Canton	5	Madison/East End	33
Cedonia/Frankford	6	Medfield/Hampden/Woodberry/Remington	34
Cherry Hill	7	Midtown	35
Chinquapin Park/Belvedere	8	Midway/Coldstream	36
Claremont/Armistead	9	Morrell Parkk/Violetville	37
Clifton-Berea	10	Mount Washington/Coldspring	38
Cross-Country/Cheswolde	11	North Baltimore/Guilford/Homeland	39
Dickeyville/Franklintown	12	Northwood	40
Dorchester/Ashburton	13	Oldtown/Middle East	41
Downtown/Seton Hill	14	Orangeville/E. Highlandtown	42
Edmonson Village	15	Patterson Park North & East	43
Fells Point	16	Penn North/Reservoir Hill	44
Forest Park/Walbrook	17	Pimlico/Arlington/Hilltop	45
Glen-Fallstaff	18	Poppleton/The Terraces/Hollins Market	46
Greater Charles Village/Barclay	19	Sandtown-Winchester/Harlem Park	47
Greater Govans	20	South Baltimore	48
Greater Mondawmin	21	Southeastern	49
Greater Roland Park/Poplar Hill	22	Southern Park Heights	50
Greater Rosemont	23	Southwest Baltimore	51
Greenmount East	24	The Waverlies	52
Hamilton	25	Upton/Druid Heights	53
Harbor East/Little Italy	26	Washington Village/Pigtown	54
Harford/Echodale	27	Westport/Mount Winans/Lakeland	55
Highlandtown	28		

Figures 3.1a-b: CSA distribution shaded by quintile of female cancer mortality (all and breast) in Baltimore City, MD per 1,000 person-years

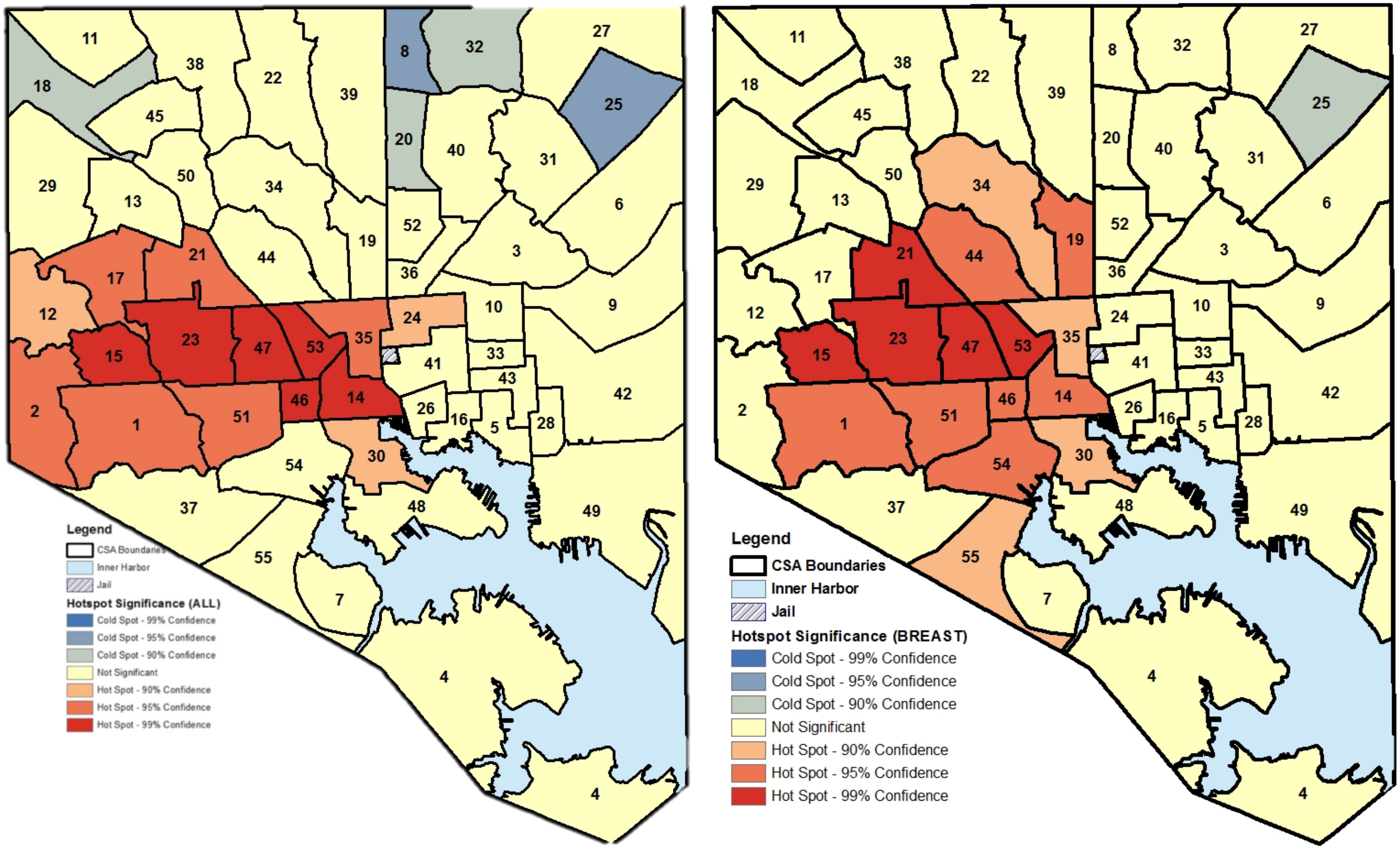


Figures 3.1c-d: CSA distribution shaded by quintile of female cancer mortality (cervical and colorectal) in Baltimore City, MD per 1,000 person-years

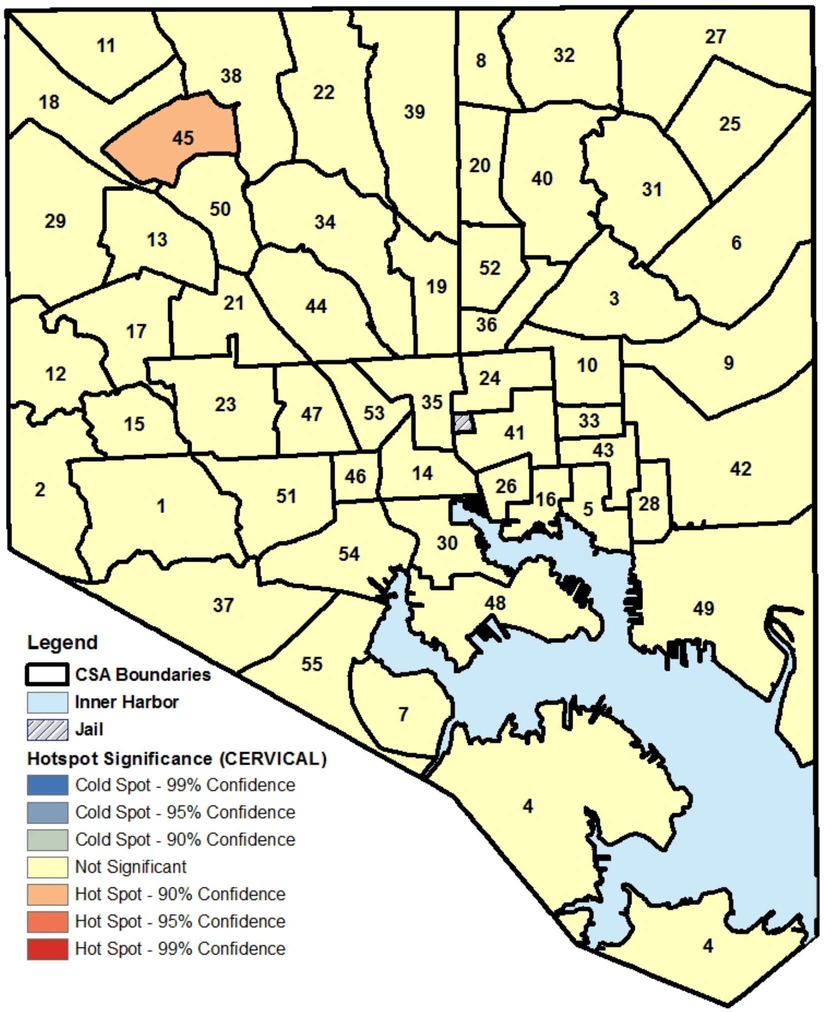




Figures 3.2a-b: Hot spot analysis and statistical significance of female cancer mortality (all and breast) per 1,000 person-years by CSA



Figures 3.2c-d: Hot spot analysis and statistical significance of female cancer mortality (cervical and colorectal) per 1,000 person-years by CSA



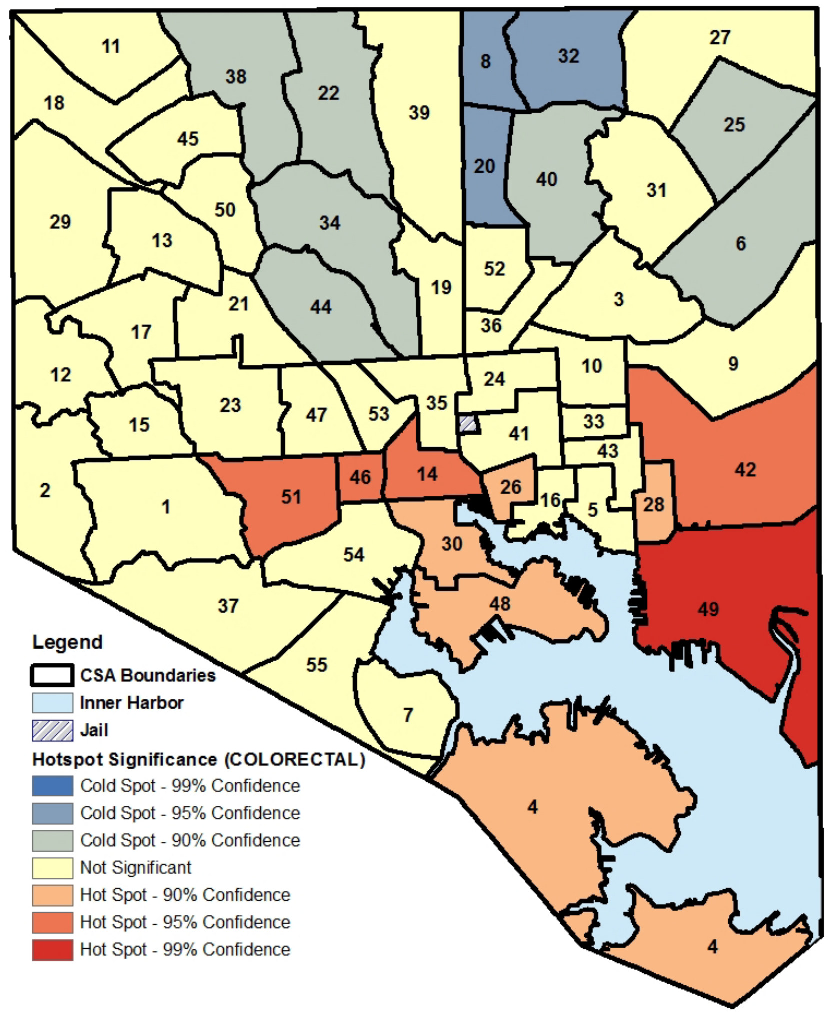


Table 3.4: Distribution of cancer site quintiles and hot/cold spatial clusters by CSA for cancer mortality

		Quintile Incidence			Hot Spot		
ID	**CSA**	**Breast**	**Cervical**	**Colorectal**	**Breast**	**Cervical**	**Colorectal**
1	Allendale/Irvington/S. Hilton						
2	Beechfield/Ten Hills/West						
3	Belair-Edison						
4	Brooklyn/Curtis Bay						
5	Canton						
6	Cedonia/Frankford						
7	Cherry Hill						
8	Chinquapin Park/Belvedere						
9	Claremont/Armistead						
10	Clifton-Berea						
11	Cross-Country/Cheswolde						
12	Dickeyville/Franklintown						
13	Dorchester/Ashburton						
14	Downtown/Seton Hill						
15	Edmonson Village						
16	Fells Point						
17	Forest Park/Walbrook						
18	Glen-Fallstaff						
19	Greater Charles Village/Barclay						
20	Greater Govans						
21	Greater Mondawmin						
22	Greater Roland Park/Poplar Hill						
23	Greater Rosemont						
24	Greenmount East						
25	Hamilton						
26	Harbor East/Little Italy						
27	Harford/Echodale						
28	Highlandtown						
29	Howard Park/West Arlington						
30	Inner Harbor/Federal Hill						
31	Lauraville						
32	Loch Raven						
33	Madison/East End						
34	Medfield/Hampden/Woodberry/Remington						
35	Midtown						
36	Midway/Coldstream						
37	Morrell Parkk/Violetville						
38	Mount Washington/Coldspring						
39	North Baltimore/Guilford/Homeland						
40	Northwood						
41	Oldtown/Middle East						
42	Orangeville/E. Highlandtown						
43	Patterson Park North & East						
44	Penn North/Reservoir Hill						
45	Pimlico/Arlington/Hilltop						
46	Poppleton/The Terraces/Hollins Market						
47	Sandtown-Winchester/Harlem Park						
48	South Baltimore						
49	Southeastern						
50	Southern Park Heights						
51	Southwest Baltimore						
52	The Waverlies						
53	Upton/Druid Heights						
54	Washington Village/Pigtown						
55	Westport/Mount Winans/Lakeland						

Quintile Key

Lowest quintile — Highest quintile

Cluster Key

Cold Spot- 99% CI — Not Significant — Hot Spot- 99% CI

Table 3.5a: Unadjusted Ordinary Least Squares regression models for cancer mortality by cancer site and candidate neighborhood-level covariates

	All Cancers		Breast		Cervical		Colorectal	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Unadjusted models								
Aged 50-74 yrs	0.001	0.901	0.001	0.979	0.046	0.036	-0.021	0.051
% African American	0.238	<0.001*	0.217	<0.001*	-1.247	1.024	-0.082	0.669
Racial Diversity Index	-0.121	0.220	-0.053	0.616	-0.551	1.700	0.031	0.920
Household income <25K	0.656	<0.001*	0.723	<0.001*	0.042	1.131	0.100	0.842
Female headed	0.515	<0.001*	0.521	<0.001*	-0.252	2.191	0.149	0.654
Vacants	0.677	<0.001*	0.753	<0.001*	7.439	10.946	-0.058	0.928
Housing violations	4.747	<0.001*	5.351	<0.001*	-0.303	0.312	-0.880	0.786
Crime	0.038	0.206	0.069	0.027*	-0.525	1.421	0.071	0.440
Domestic violence	0.647	<0.001*	0.576	<0.001*	-0.137	0.654	0.630	0.129
Teen births	0.300	<0.001*	0.393	0.037*	-1.452	1.898	0.393	0.037*
Employed	-0.658	<0.001*	-0.680	<0.001*	-0.040	0.040	0.348	0.535
Businesses	-0.001	0.854	0.004	0.291	1.182	2.389	-0.005	0.663
Voted	-0.751	<0.001*	-0.703	0.003	0.109	0.370	-0.419	0.552
Dirty streets	0.102	0.003*	0.048	0.204	1.929	1.248	0.164	0.128
Tree coverage	-0.295	0.013*	-0.226	0.078	0.155	3.998	-0.622	0.094
Neighborhood associations	0.794	0.034*	0.668	0.097	0.046	0.036	-0.026	0.982

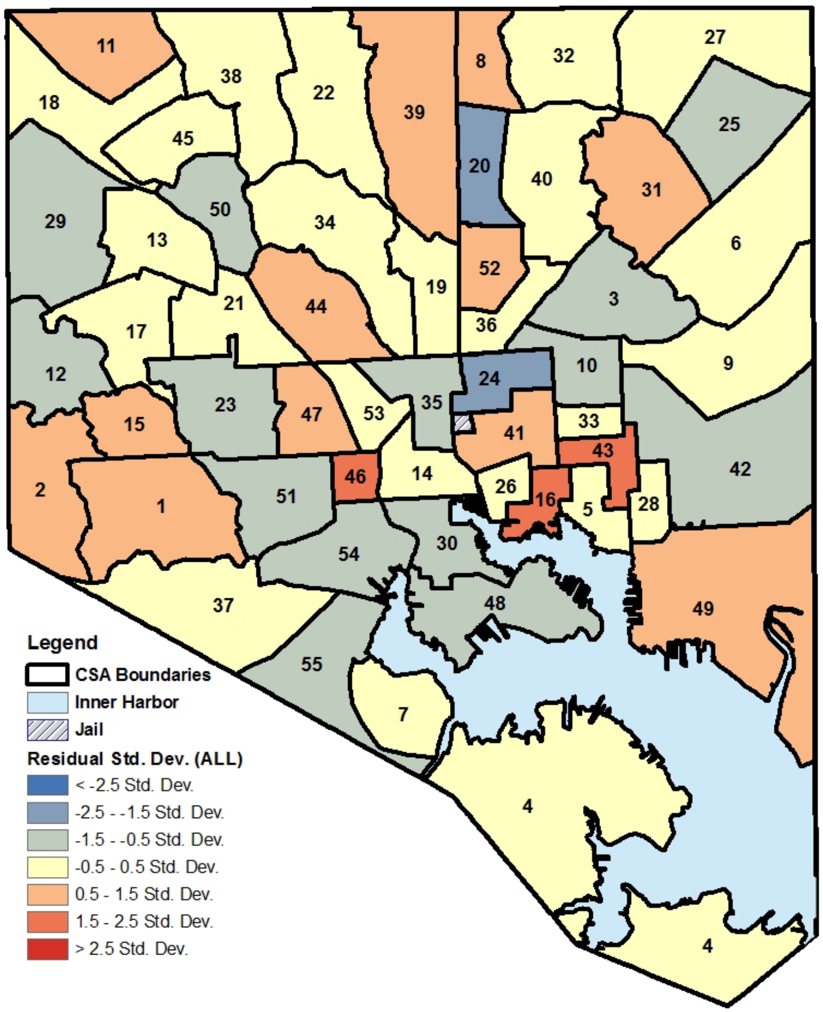
* Statistically significant

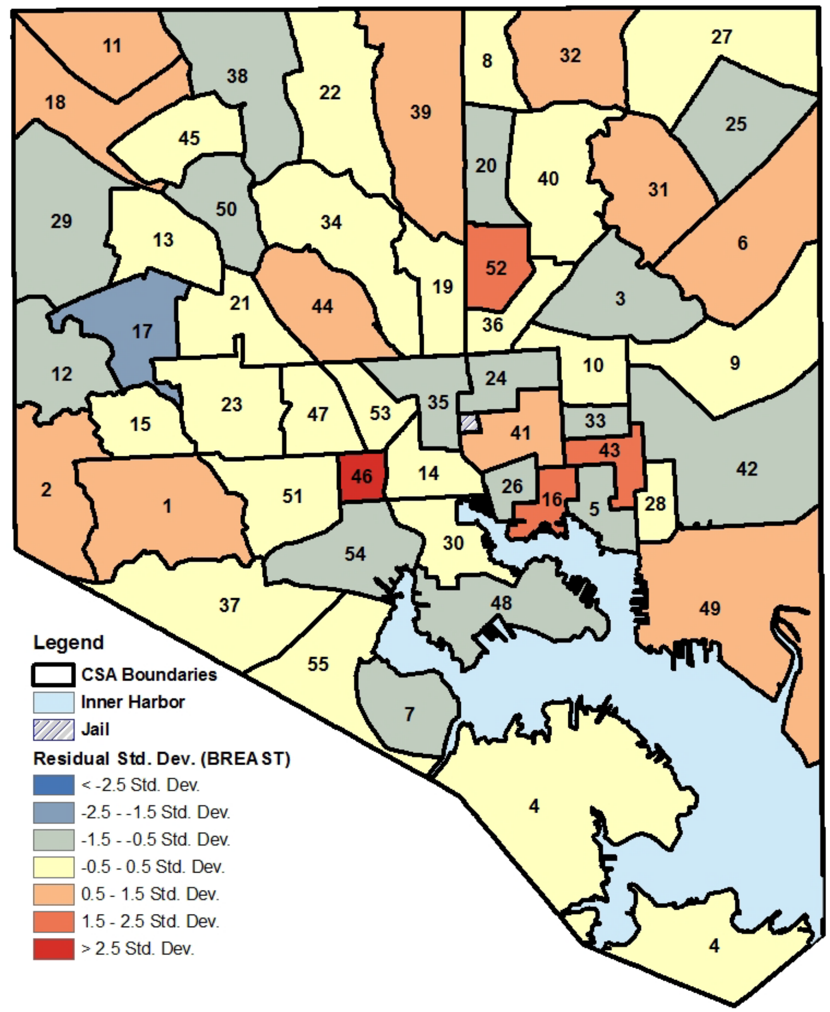
Table 3.5b: Adjusted Ordinary Least Squares regression models for cancer mortality by cancer site and candidate neighborhood-level covariates

	All Cancers		Breast		Cervical**		Colorectal	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Adjusted models								
Aged 50-74 yrs	0.002	0.472	0.001	0.979				
% African American			-0.118	0.252				
Household income <25K			-0.004	0.987				
Female headed	0.295	0.112	0.492	0.063				
Vacants	-0.423	0.141	-0.166	0.573				
Housing violations	3.766	0.004*	5.281	0.003*				
Crime			0.027	0.353				
Domestic violence	0.141	0.575	0.157	0.594				
Teen births	0.148	0.061	0.023	0.801			0.393	0.037*
Employed	0.389	0.141	0.423	0.157				
Businesses								
Voted	0.214	0.385	0.296	0.350				
Dirty streets	0.029	0.386						
Tree coverage	-0.100	0.370						
Neighborhood associations	0.277	0.400						
R-squared	0.561		0.461				0.062	

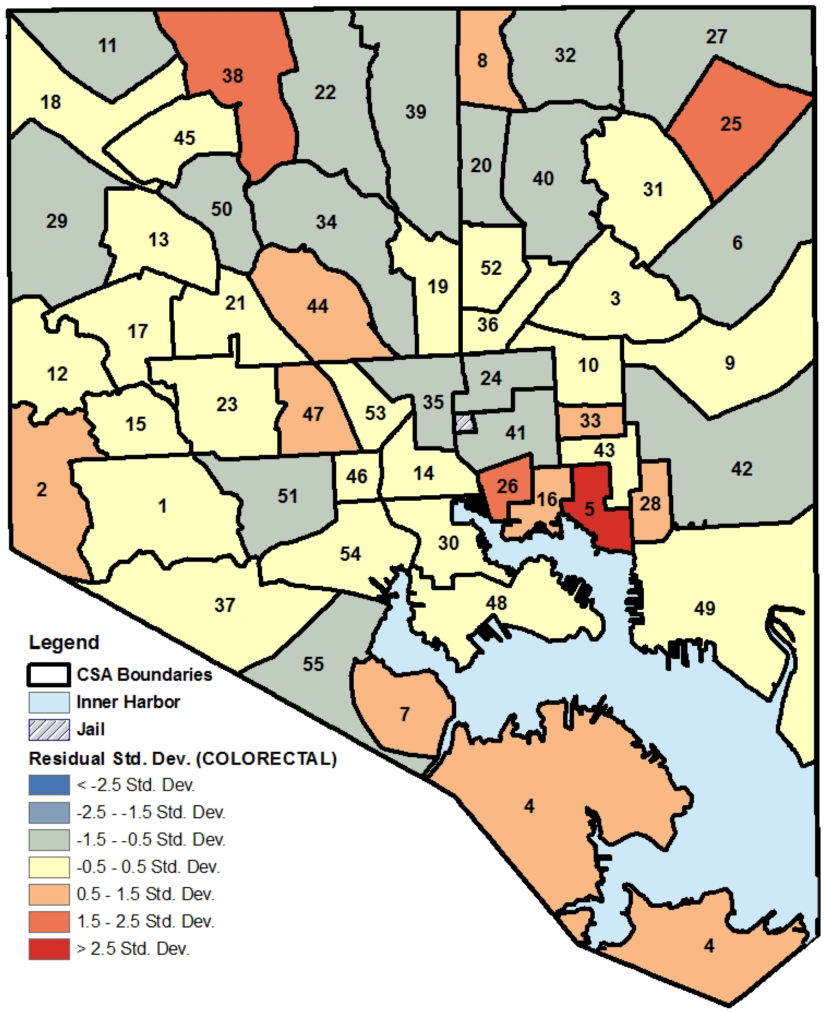
* Statistically significant
** Did not yield significant unadjusted associations

Figures 3.3a-b: Spatial output of final models for ordinary least squares regression for cancer mortality (all and breast)

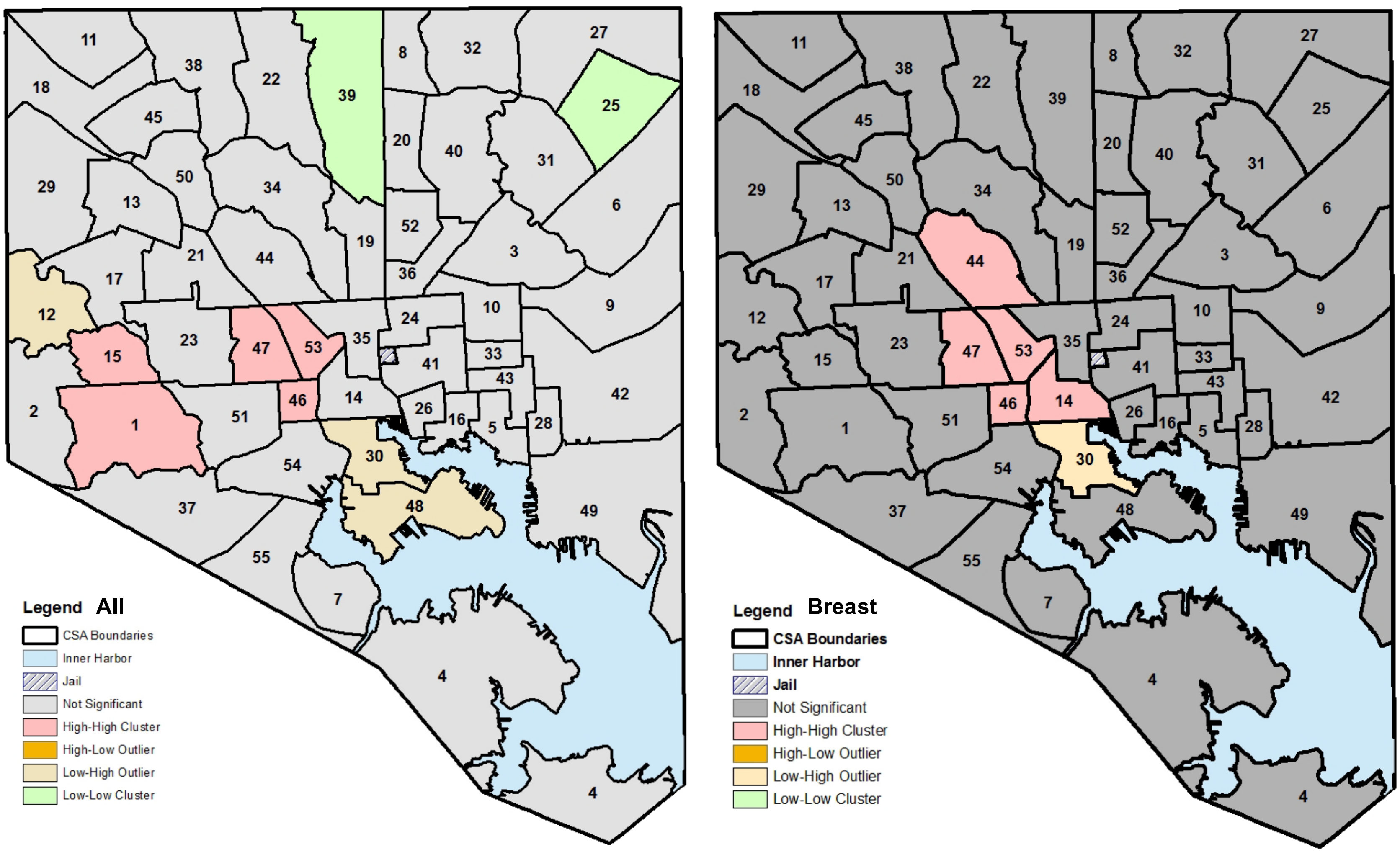




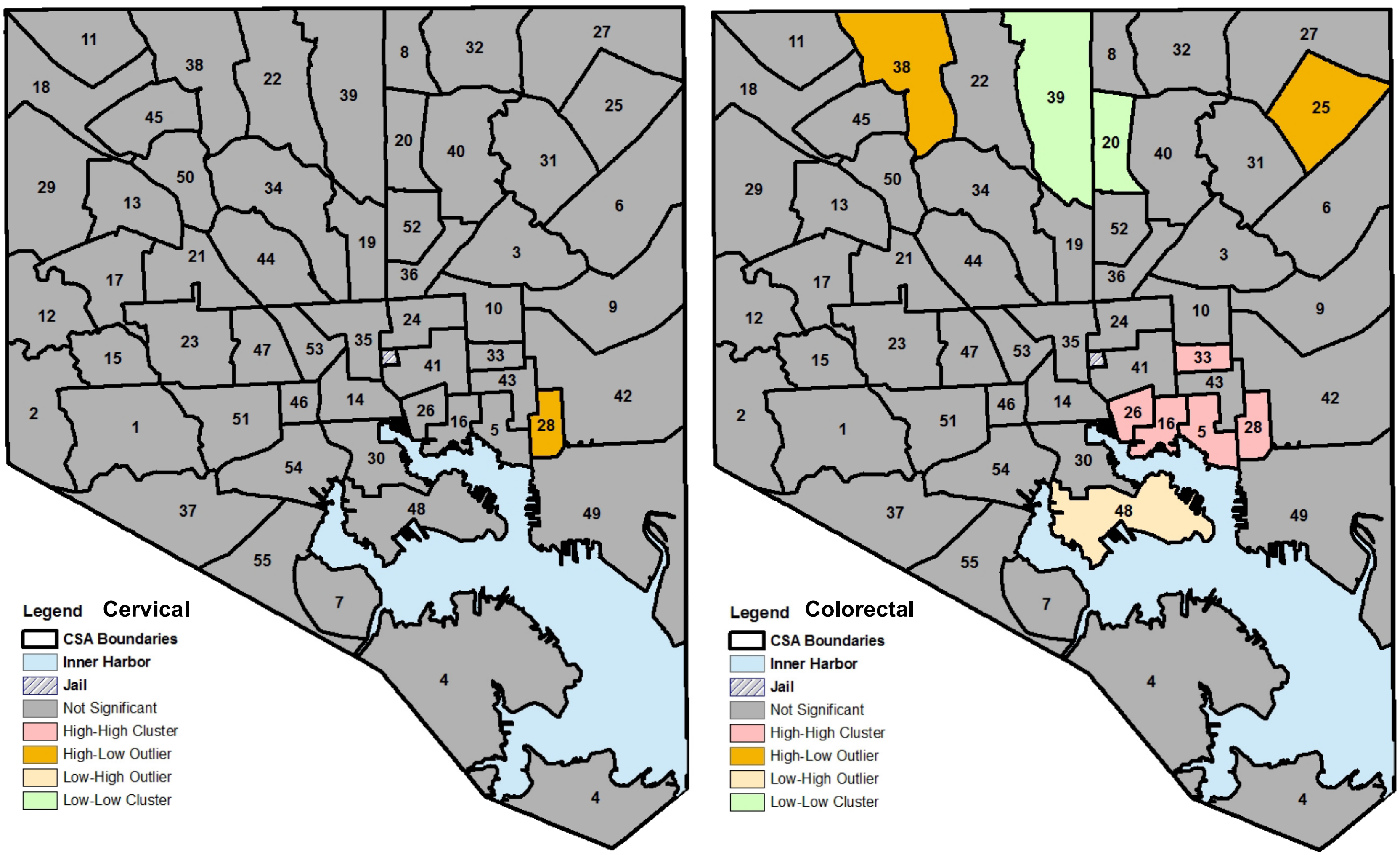
Figures 3.3c: Spatial output of final models for ordinary least squares regression for cancer mortality (colorectal)



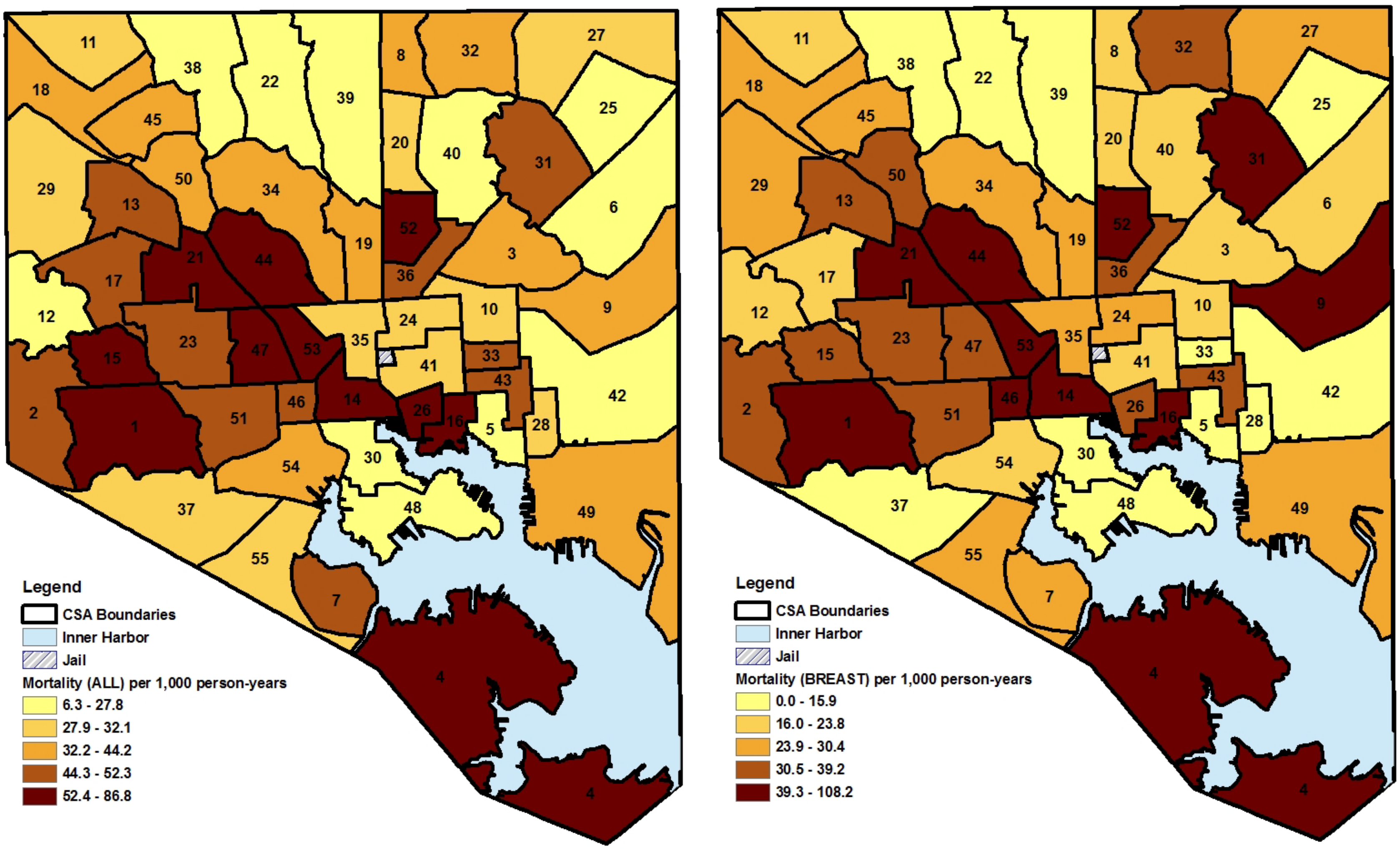
Figures 3.4a-b: Local Moran's I analysis for cancer mortality (all and breast)



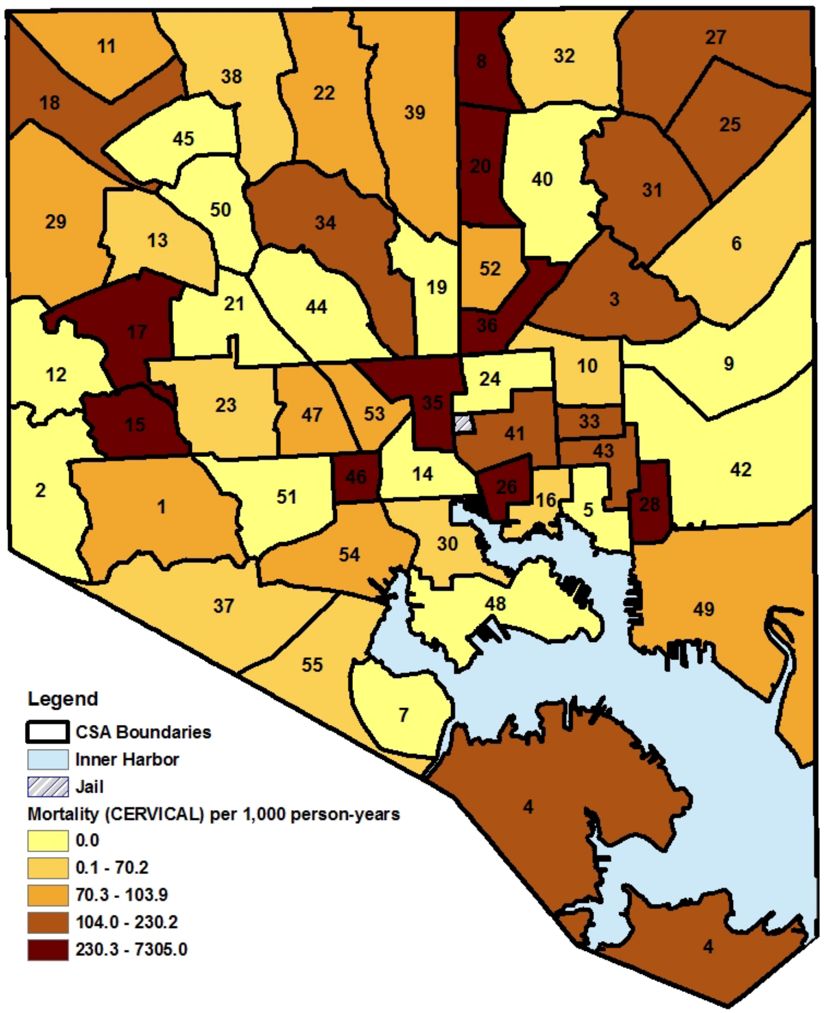
Figures 3.4c-d: Local Moran's I analysis for cancer mortality (cervical and colorectal)

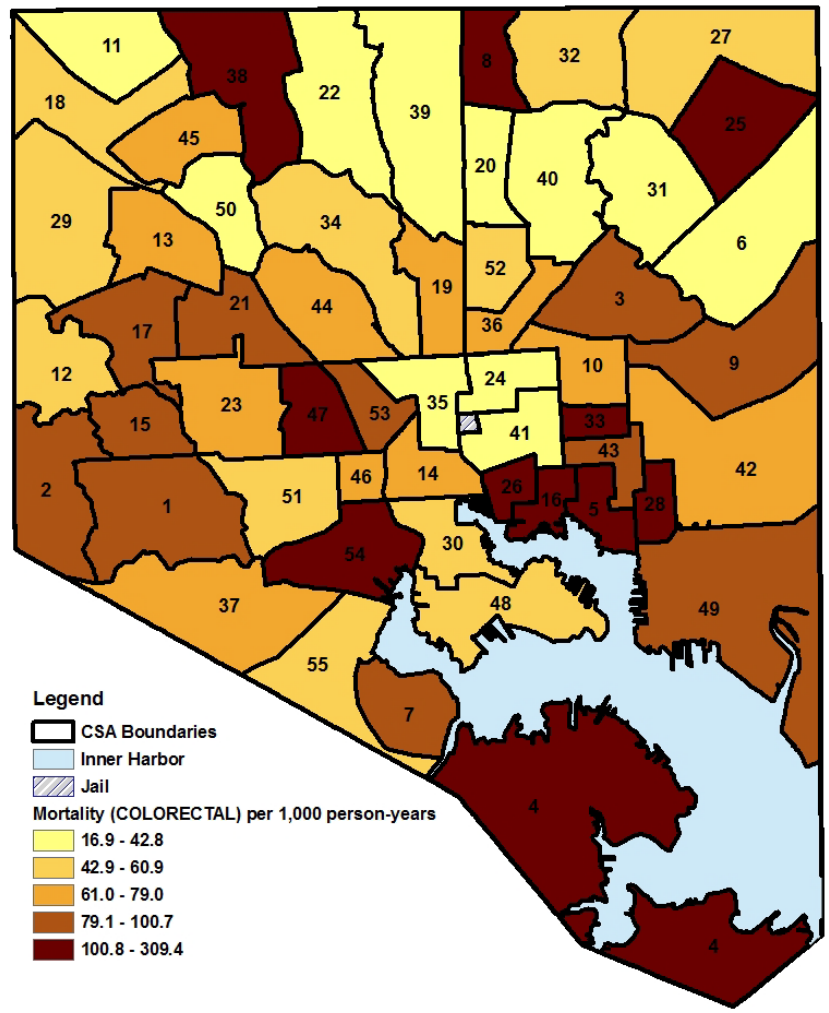


Appendix 3A Figure S3.1a-b: CSA distribution shaded by quintile of female cancer mortality (all and breast) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years

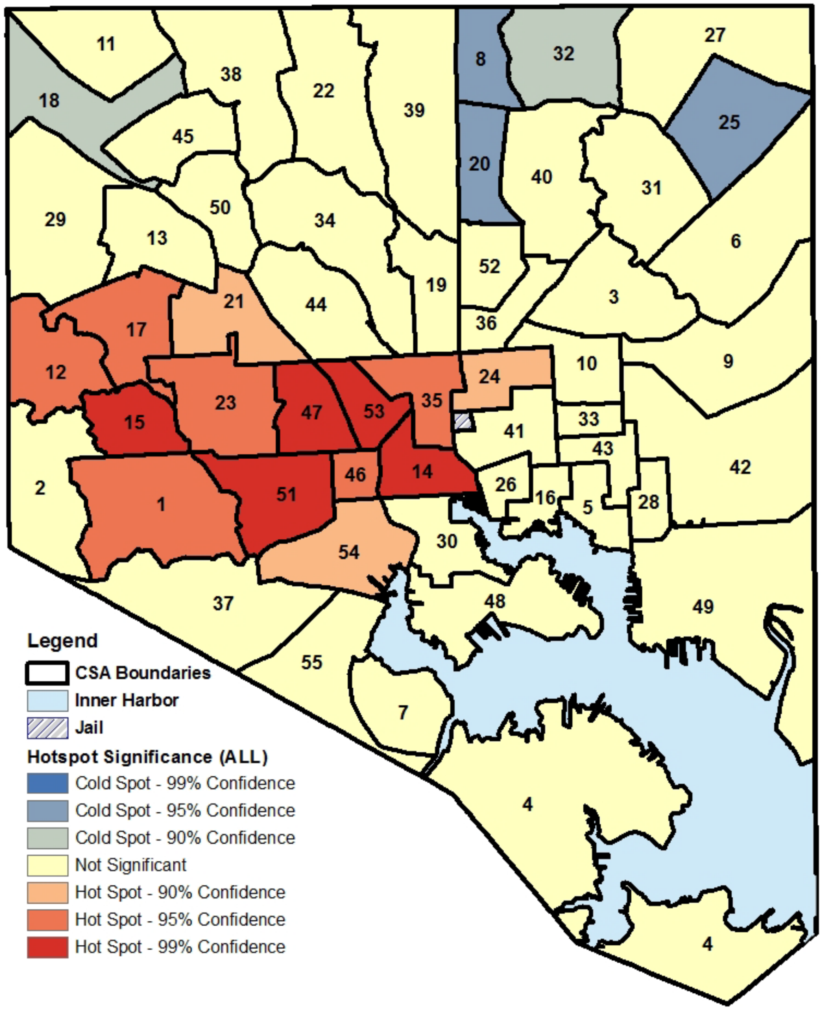


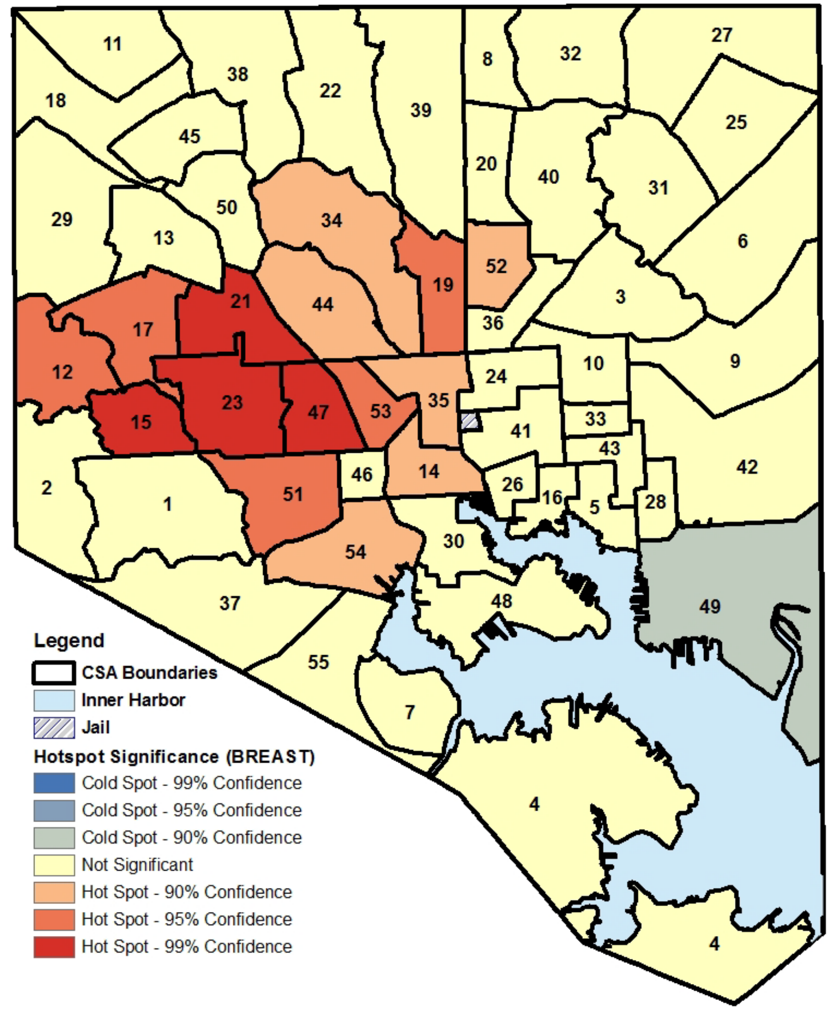
Appendix 3A Figure S3.1c-d: CSA distribution shaded by quintile of female cancer mortality (cervical and colorectal) in Baltimore City, MD per 1,000 female residents aged 50 to 74 years



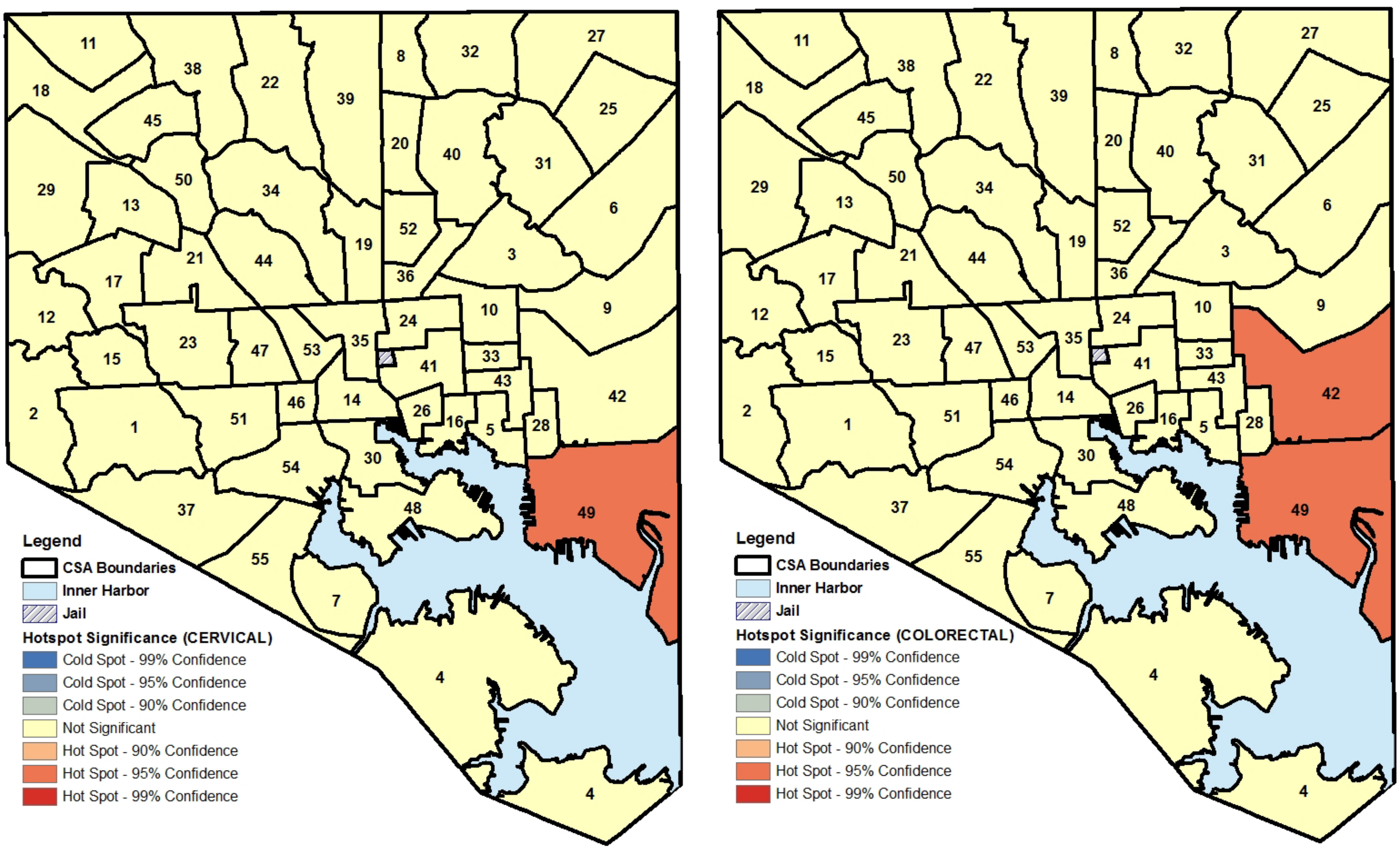


Appendix 3A Figure S3.2a-b: Hot spot analysis and statistical significance of cancer mortality (all and breast) per 1,000 female residents aged 50 to 74 years by CSA





Appendix 3A Figure S3.2c-d: Hot spot analysis and statistical significance of cancer mortality (cervical and colorectal) per 1,000 female residents aged 50 to 74 years by CSA



Appendix 3A Table S3.1a: Unadjusted Ordinary Least Squares regression models for cancer mortality among females 50-74 years by site and candidate neighborhood-level covariates

	All Cancers		**Breast**		**Cervical**		**Colorectal**	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Unadjusted models								
Aged 50-74 yrs	0.001	0.952	0.002	0.645	-0.070	0.778	-0.025	0.060
% African American	0.207	0.002*	0.221	0.003*	-0.477	0.913	-0.290	0.225
Racial Diversity Index	-0.093	0.398	-0.049	0.686	4.403	0.525	0.264	0.491
Household income <25K	0.522	0.003*	0.604	0.002*	-7.256	0.523	-0.428	0.495
Female headed	0.482	<0.001*	0.478	<0.001*	-5.839	0.438	-0.281	0.501
Vacants	0.512	0.025*	0.562	0.026*	-8.632	0.555	-0.358	0.659
Housing violations	4.627	<0.001*	5.528	<0.001*	-50.349	0.492	-5.000	0.216
Crime	0.074	0.023*	0.082	0.023*	0.696	0.741	0.025	0.833
Domestic violence	0.641	<0.001*	0.560	<0.001*	-2.432	0.798	0.158	0.764
Teen births	0.298	<0.001*	0.206	0.005*	0.380	0.931	0.129	0.592
Employed	-0.572	0.003*	-0.597	0.006*	13.707	0.280	1.136	0.103
Businesses	0.004	0.373	0.006	0.182	-0.033	0.904	-0.009	0.570
Voted	-0.781	0.001*	-0.712	0.009*	14.140	0.375	0.845	0.338
Dirty streets	0.074	0.055	0.022	0.621	1.841	0.455	0.162	0.234
Tree coverage	-0.225	0.093	-0.068	0.652	-9.176	0.279	-0.113	0.811
Neighborhood associations	0.753	0.072	0.874	0.059	-26.360	0.322	-1.175	0.426

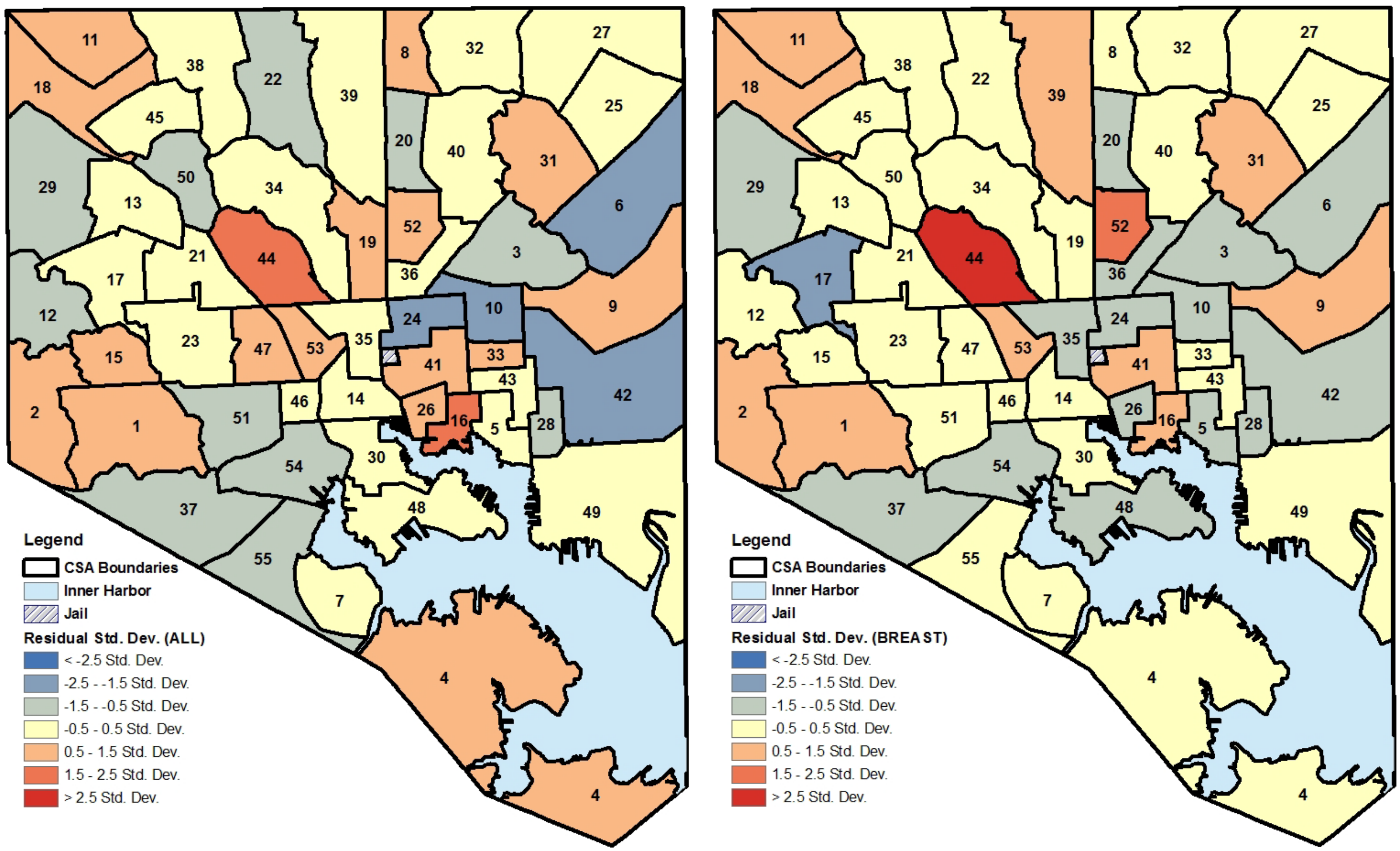
* Statistically significant

Appendix 3A Table S3.1b: Adjusted Ordinary Least Squares regression models for cancer mortality among females 50-74 years by site and candidate neighborhood-level covariates

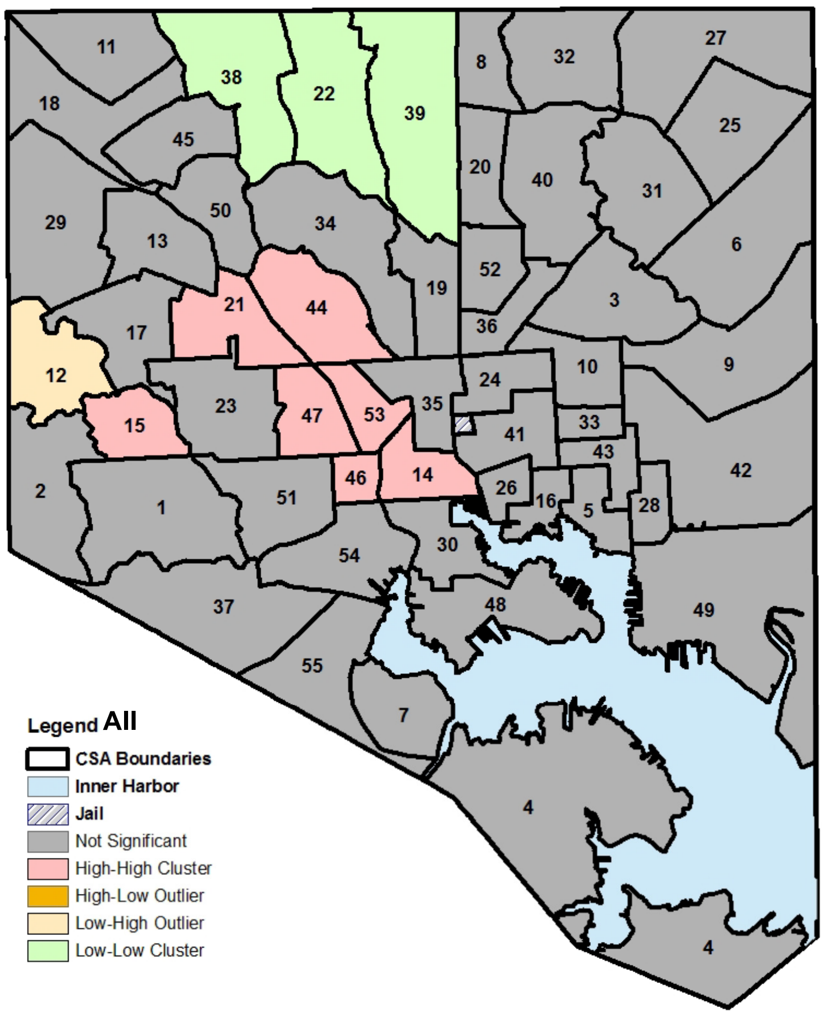
	All Cancers		Breast		Cervical**		Colorectal**	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Adjusted models								
Aged 50-74 yrs	0.002	0.509	0.005	0.232				
% African American	0.015	0.854	-0.086	0.496				
Racial Diversity Index								
Household income <25K	-0.459	0.110	-0.374	0.280				
Female headed			0.465	0.157				
Vacants	-0.546	0.068	-0.715	0.055				
Housing violations	6.292	<0.001*	8.085	<0.001*				
Crime	0.044	0.162	0.048	0.204				
Domestic violence	0.330	0.237	0.193	0.594				
Teen births	0.198	0.045*	0.044	0.704				
Employed	0.115	0.690	0.379	0.295				
Voted	0.091	0.778	0.128	0.745				
R-squared	0.498		0.402					

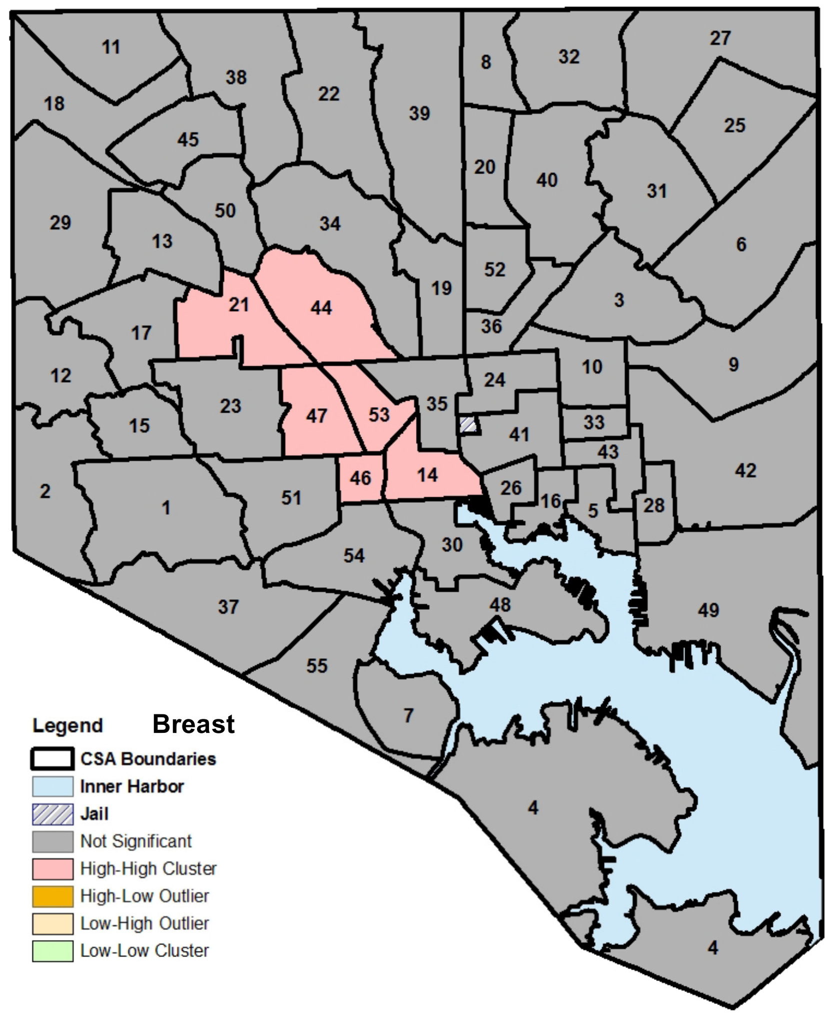
* Statistically significant
** Did not yield significant unadjusted associations

Appendix 3A Figure S3.3a-b: Spatial output of final models for ordinary least squares regression for cancer mortality (all and breast) of females aged 50 to 74 years

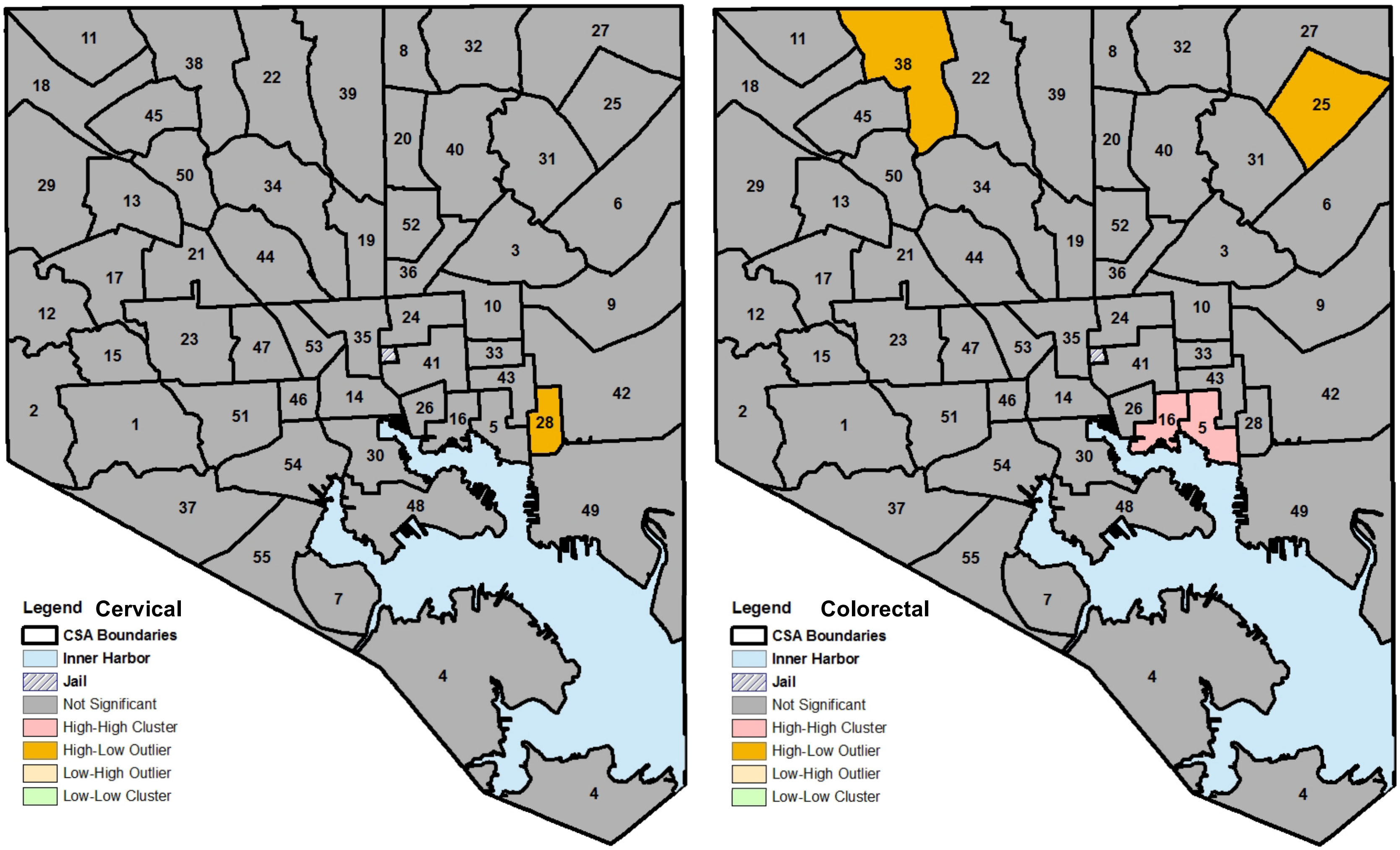


Appendix 3A Figure S3.4a-b: Local Moran's I analysis for cancer mortality (all and breast) in females aged 50 to 74 years

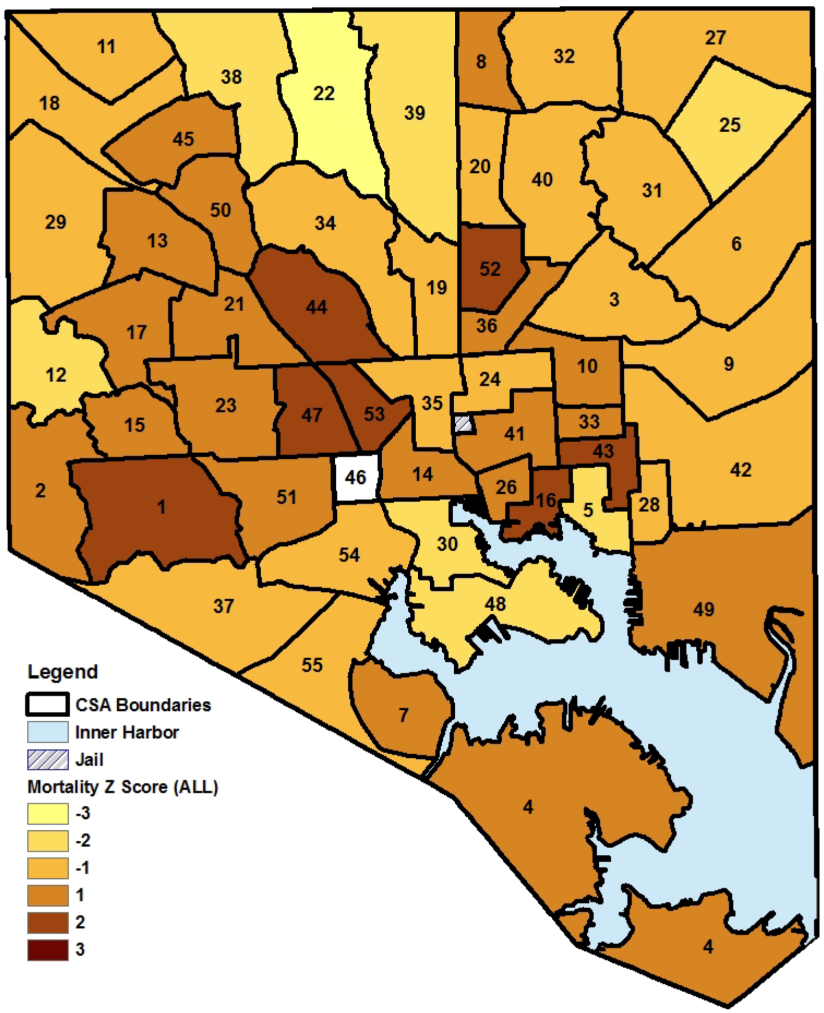


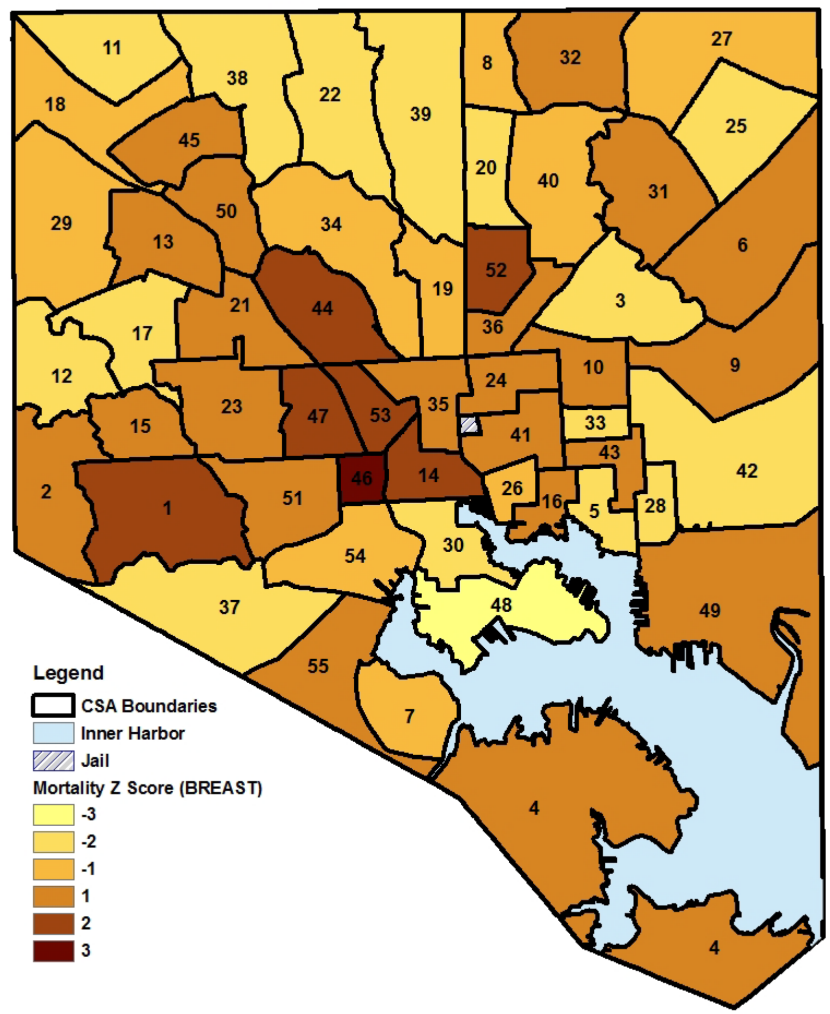


Appendix 3A Figure S3.4c-d: Local Moran's I analysis for cancer mortality (cervical and colorectal) in females aged 50 to 74 years

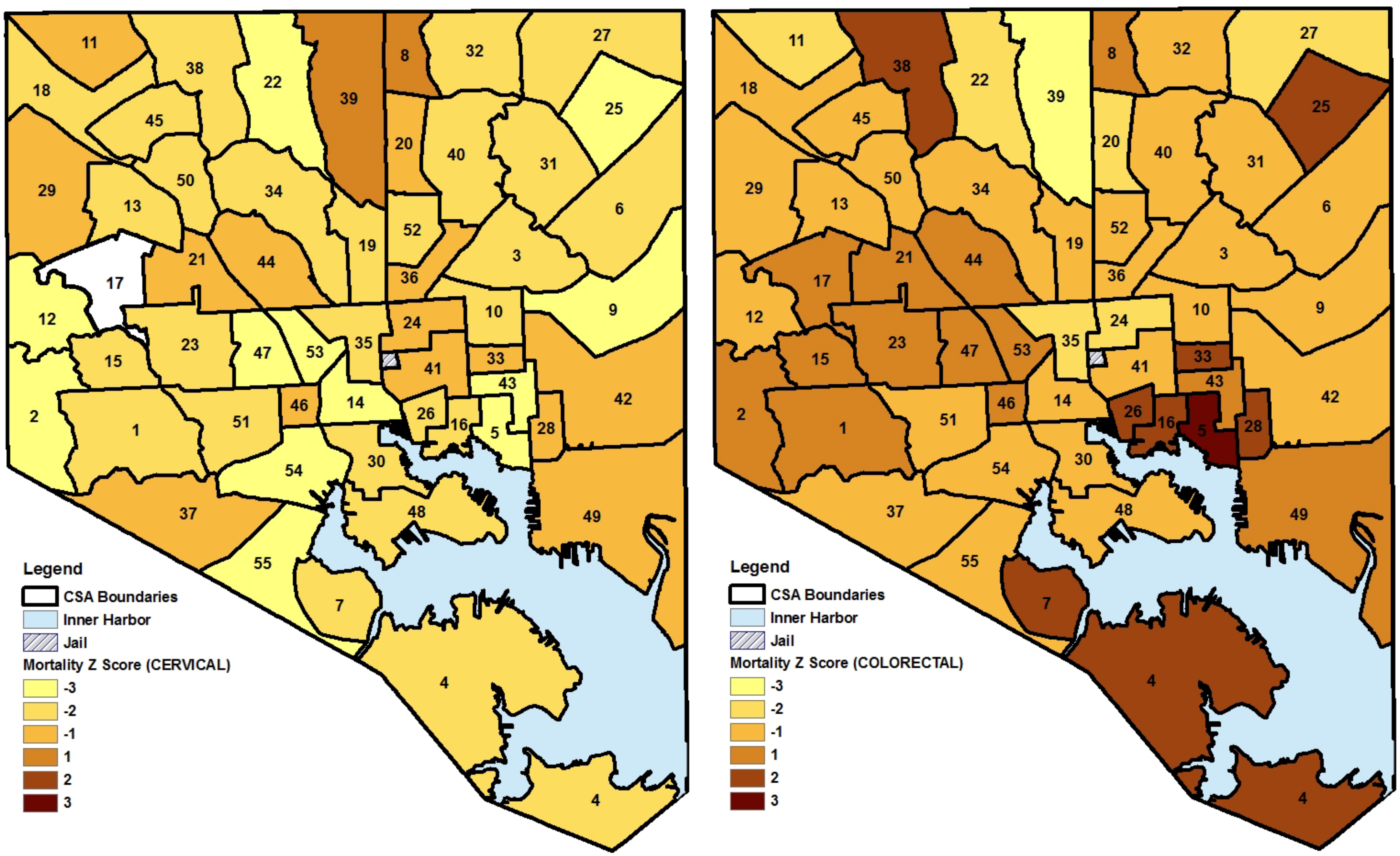


Appendix 3B Figure S3.1a-b: CSA distribution shaded by Z-score of female cancer deaths (all and breast) in Baltimore City, MD per 1,000 person-years





Appendix 3B Figure S3.1c-d: CSA distribution shaded by Z-score of female cancer deaths (cervical and colorectal) in Baltimore City, MD per 1,000 person-years



CHAPTER 4: EVALUATING NEIGHBORHOOD CORRELATES AND GEOGRAPHIC DISTRIBUTION OF CANCER TREATMENT AND SCREENING FACILITIES

Abstract
Objective: To evaluate the geographic variation of breast and colorectal screening/treatment facilities in Baltimore City and the neighborhood characteristics associated with the observed variation.

Design and Methods: Using publicly available databases and directories, Baltimore City facilities providing services (i.e., screening and treatment) for breast and colorectal cancer were identified. In order to be included in the analysis, the facility needed to meet two main criteria. Due to the focus on current access to these resources, facilities needed to have been open or in service as of January 1, 2015. Additionally, as a proxy for high-quality care, only centers that were recognized by an accreditation body were included. Physical access to these facilities was defined in two ways for each neighborhood:
1) cancer service density (organizations per square mile); and 2) cancer service rate (organizations per 10,000 female residents aged 50 to 74 years). Using these measures as the outcome, geographical variability was evaluated through thematic maps, hot/cold spot identification, and the local Moran's I. The neighborhood characteristics associated with the geographic variation were evaluated through ordinary least squares regression models.

Results: A total of 17 centers were identified within the boundaries of Baltimore City. Cancer facilities for breast and colorectal cancer services were primarily located in the downtown area of Baltimore City. The rest of the city had little to no physical access to cancer resources. When evaluating which of the community-level covariates explained this geospatial variability, crime and total number of businesses were significantly

associated with cancer service density while cancer service rate was associated with crime, number of businesses, and the Racial Diversity Index. These adjusted models explained the majority of the geographic distribution as indicated by each model's R-squared statistic of 71.6% for service rate and 81.7% for service density.

Conclusion: Consideration should be given to first identifying the current allocation of health resources and then deciding if placement is warranted based on the location's disease burden. Also, information on screening/treatment locations should be more readily available and easily accessible in a central place to help better inform community members seeking care.

Introduction

From the previous chapters, several of the results have demonstrated significant geographic variation in cancer outcomes within Baltimore City. While the distribution of cancer-related events, such as incidence and mortality, are important to highlight, the location of cancer resources and facilities should not be overlooked. As the prominence of cancer disparities continues to come to the forefront, there is a noticeable gap in effectively leveraging innovation across all populations. "There is a critical disconnect between what we discover and what we deliver, between what we know and what we do for all people."[1]

There has been some research, although limited, that evaluates the local distribution of healthcare resources, such as cancer facilities. Using the assumption that proximity is related to access, spatial methodology has been utilized to conduct asset mapping and describe the physical network of prevention as well as treatment services. Most of the literature in this specific area has focused primarily on mammography screening with some discussion on colorectal cancer resources.[2–8] Although these efforts have sought to describe the physical presence of healthcare services in small areas, there has been less work conducted to identify and understand potential neighborhood characteristics associated with the observed geographic distribution.

The need to understand the geographical availability of healthcare services has motivated not only research; it has also sparked policy changes within Baltimore City. In 2012, the Maryland Health Improvement and Disparities Reduction Act was passed to address geographic disparities across health outcomes by designating Health Enterprise Zones in neighborhoods that lacked resources.[9]

This chapter sets out to evaluate the geographic distribution of cancer facilities in Baltimore City, Maryland as well as the neighborhood characteristics associated with the cancer service rate and density observed. Spatial statistics will be the primary analytical method in order to allow for the utilization of area-level covariates that describe the social determinants of health across Baltimore City communities in a much more comprehensive way than the often used U.S. Census indicators. The aim of this chapter is to explore whether areas with more cancer resources tend to have similar neighborhood characteristics.

Methods

Cancer facility ascertainment

Baltimore City directories and publicly available online databases served as the sources for identifying the locations of relevant cancer-related services for breast and colorectal cancer in Baltimore City. The listing of facilities consists of places that were open as of January 1, 2015. In order to be included in the analysis, the sites rendering the services had to be recognized by an accreditation body. The Mammography Facility Database was utilized to obtain the addresses for sites that provided this breast screening service. This database is updated periodically by the FDA based on information received from the American College of Radiology, which is an FDA-approved accreditation body.[10] The Commission on Cancer (CoC) of the American College of Surgeons was also utilized to identify cancer hospitals that provided services related to breast or colorectal cancer. Since over 70% of patients in the United States receive cancer treatment from a facility accredited by the CoC, this database contained the sites within Baltimore City where residents would mostly likely obtain care upon receiving a cancer diagnosis.

For outpatient sites offering colonoscopies outside of the hospital setting, accreditation was required from the Accreditation Association for Ambulatory Health Care (AAAHC), American Association for Accreditation of Ambulatory Surgery Facilities (AAAASF), or the Joint Commission on Accreditation for Healthcare Organizations (JCAHO). The Maryland Ambulatory Surgery Provider Directory, which noted the accreditation of all ambulatory surgery centers, served as the primary source for outpatient colonoscopy facilities.[11] Accreditation was required for inclusion to serve as a proxy for identifying a facility that was of high quality.

Neighborhood characteristics

As previously mentioned in Chapters 2 and 3, neighborhood-level characteristics from the Vital Signs report provided the primary independent variables for Baltimore City using the Community Statistical Area (CSA) as the geographic unit for analysis.[12] The methodology of assessing the correlation among these community covariates is the same as those described in those chapters. Each area-level characteristic was summarized as the mean value across the years of data available based on the stability seen in the trend lines plotted in Chapter 2.

Statistical analyses

The prevention and treatment sites were aggregated by CSA to construct measures capturing physical access to cancer screening and care. These measures consisted of cancer service density (organizations per square mile) and cancer service rate (organizations per 10,000 female residents aged 50 to 74 years). In certain cases, facilities within the same healthcare network would delegate services to multiple sites within walking distance of each other. In these instances, this was counted as one facility since the disaggregation of various services across proximal locations was considered comparable to a single site offering all of the various services within the same building. Choropleth maps were produced to provide an exploratory visualization of the distribution by quintile of cancer service density and cancer service rate across the CSAs in Baltimore City. The spatial aggregation of cancer service observed in these initial thematic maps was verified using the Getis-Ord Gi* statistic. This spatial cluster detection method identified CSAs that had either significantly more or less than expected cancer service as compared to Baltimore City as a whole.[13]

Upon discovering statistically significant spatial clusters, global ordinary least squares (OLS) regression was conducted for each cancer site to identify which of the neighborhood characteristics was associated with and could explain the geographic distribution of cancer service. These models yielded an R-statistic, which indicated what percentage of the observed geographic variability could be explained by the CSA-level covariates included in the model.

As an additional cluster detection method, a local indicators of spatial association (LISA) function using Local Moran's I was also integrated into the analysis to determine if there were any discordant clusters pertaining to physical access in Baltimore City. The strength of a LISA analysis is that it illustrates not only the degree to which each CSA affects the global magnitude but it also brings attention to significant cluster outliers.[14,15] The identification of cluster outliers is particularly notable since it has the potential to highlight neighboring CSAs for more in-depth research to better understand their significant difference in outcomes despite their immediate proximity. The spatial statistics and mapping visualizations described above were carried out with ArcGIS 10.3 software.[16]

All data utilized for these analyses were obtained through publicly available resources and datasets.

Results

In total, there were 17 accredited cancer-related facilities that provided a form of cancer screening or treatment. The point pattern map showing the locations of these facilities (**Figure 4.1**) illustrated that the majority of these sites were located in the downtown area of Baltimore City. There were only 4 sites providing coverage to the northern sections of county. Each facility provided a type of cancer-related service, in the form of either screening or treatment for any of the three primary sites of interest (**Table 4.1**). With only a few exceptions, such as the radiology service, most of the sites conducted services that spanned across the disease spectrum (i.e., carrying out both screening and treatment). To help navigate through the subsequent maps, the CSA map key has been provided again (**Table 4.2**).

The choropleth maps for cancer service density and service rate were reflective of what was observed in the initial point pattern map (**Figures 4.2a-b**). The CSA of Downtown/Seton Hill had the most facilities, which clearly elevated both the service density and service rate measures. The entire northeastern portion of Baltimore City had little physical access to facilities. Unsurprisingly, the hot spot analysis primarily brought attention to the downtown area of Baltimore City. This portion of the city had higher a density and rate of service as compared to the overall measure for Baltimore City as a whole (**Figures 4.3a-b**). Upon running the OLS model for both measures, crime and total number of businesses were significant for service density while service rate had these same significant indicators as well as the Racial Diversity Index (**Table 4.3)**.

Much like the situation observed in Chapter 2, the directionality of the association between crime and physical access was unexpected. The coefficients indicated that increased crime was associated with increased access. Also mentioned in Chapter 2, this relationship might be due to population size. Crime tends to occur in more urbanized and densely populated areas while screening and treatment centers may also be placed in similarly urbanized and populated areas to have a larger catchment network. The final spatial analysis conducted, the local Moran's I, reiterated what was already observed through the previous cluster detection methods (**Figures 2.5a-b**). There was no evidence of any discordant outliers. The clustering was limited to only high-high neighborhoods (i.e., CSAs with high access that neighbor other CSAs with high access) exclusively in the downtown area by Inner Harbor.

Conclusion

The general findings of this chapter were: 1) the observed geographic variation in the locations of cancer facilities within Baltimore City; 2) the similarity seen in the distribution of cancer service rate and cancer service density; and 3) the similarity in the neighborhood characteristics that were associated with both of the constructed measures for cancer service access. The overall results aligned with the a priori hypothesis, which anticipated a clustering of cancer resources within Baltimore City. More specifically, the concentration of facilities was higher, relative to both population and CSA size, in the downtown portion of the city as well as Inner Harbor.

This exploratory analysis provides an early indication that groups of Baltimore City female residents have less healthcare access, as defined by physical proximity to resources. As the geographic variation shows, there appears to be an opportunity to further evaluate, and likely reconsider, the current allocation of resources, especially when factoring in the geographic variation of disease outcomes. At initial glance, it appears as though the northeastern portion of Baltimore City lacks a noticeable health resource presence. Unlike previous chapters where the significantly associated CSA-level characteristics varied across outcomes, crime and the number of businesses in a community explained both service rate and service density. This would demonstrate that the mechanisms for each of the two measures within a neighborhood context are likely very similar.

This chapter's methodology has several strengths and limitations. One of the first things that should be noted is that the service measures were calculated based on each CSA's female population and square mileage. However, the analysis plan was unable to account for whether the residents of a specific CSA actually utilized the services immediately available in their residential neighborhood. While proximity would address some of the physical barriers pertaining to access, residents might use other facilities due to a number of reasons. For example, individuals may choose to get screening or treatment services near their workplaces to better accommodate appointments or may use other doctors that might be more compatible with their specific insurance plans. As a result, cancer service may not accurately portray complete healthcare access for residents of a CSA. Future analyses would likely require a form of claims linkage to determine where individuals are actually receiving their cancer screenings or treatment for a cancer diagnosis.

An additional caveat of the analysis is that only cancer resources within Baltimore City were included in the spatial assessment. This is relevant because although communities on the outer edges of the city's boundaries may appear to lack physical access, there might be facilities in Baltimore County that are closer in proximity. Baltimore County was excluded due to: 1) the lack of a comparable geographic unit to the CSA boundaries; and 2) the absence of area-level characteristics that had the breadth and depth as those captured by the Baltimore Neighborhood Indicators Alliance. This again demonstrates the importance of other cities and counties needing to monitor measures, particularly for small areas, that comprehensively capture the social and physical conditions that can play a role in health outcomes.

Furthermore, it is also possible that zoning laws designating what can and cannot be built within a CSA might cause some areas to appear as though they have low service density. There are portions of CSAs that are required to be predominantly residential, which would make any intervention centered on improving proximal access to health resources slightly more difficult. In these instances, it could very well be that a public health need has been identified but require a solution beyond simply a health facility structure due to land use regulations. This would require either a revisiting of the zoning laws, specifically in terms of health facilities, or other alternatives to improve access, such as increased utilization of mobile units.

The analysis also limited its focus to only breast and colorectal cancer facilities. Cervical cancer resources were excluded due to the complexity of accurately and completely identifying all locations that offered cervical screening, such as pap smears. This procedure can be done in nearly any health services setting. As a result, the neighborhood characteristics associated with a service that was not exclusively focused on cancer might yield different associations as compared to just honing in on strictly cancer-related facilities. The exclusion of cervical cancer from this analysis was primarily driven by the fact that access to a cervical cancer screening resource, such as STD clinic, may not be truly indicative of a neighborhood's access across the continuum of care. Given the broader settings in which pap smears are conducted, there is a greater likelihood of patients failing to transition to cancer treatment after an initial diagnosis or suspicion of disease in a non-oncology setting. There are likely numerous barriers in transitioning to timely and appropriate cancer care when screening and treatment are done in separate settings. This issue would have gone beyond the scope of this dissertation and would require a separate, in-depth analysis to account for this transition.

Despite these limitations, this analysis also had its share of strengths. It offered a comprehensive listing of facilities for breast and colorectal cancer in Baltimore City. Surprisingly, this publicly available data required accessing several different sources. There was no single database that provided information as to the services rendered and their respective locations. The lack of an easily searchable and accessible database could prove to be a barrier for residents of Baltimore City that are trying to efficiently identify convenient sources of care.

The methodology also offers innovation within the realm of evaluating geographic disparities. There has been limited literature on the distribution of healthcare resources, particularly for cancer, and its association with neighborhood characteristics. This perspective provides insight as to whether there are particular neighborhood types with certain physical and social profiles that tend to have a lower concentration of healthcare services. This analysis took the unique approach of studying the social environment in which healthcare facilities are located.

In certain scenarios, resources are haphazardly placed in areas without understanding the community's characteristics or disease burden. This information would be important to leverage in conjunction with the concentration of disease outcomes to ascertain whether facilities are appropriately located in areas with the most need, especially since the previous chapters demonstrated the association of contextual factors with cancer outcomes. This hinges on the added strength of the existence of CSAs within Baltimore City. These geographic boundaries allow for healthcare resources to be assessed with small-area estimates while using the social determinant measures collected by the Baltimore Neighborhood Indicators Alliance over the last decade.

References:

1. Freeman HP. Poverty, culture, and social injustice: determinants of cancer disparities. CA Cancer J Clin 2004;54(2):72–7.

2. Lian M, Jeffe DB, Schootman M. Racial and geographic differences in mammography screening in St. Louis City: a multilevel study. J Urban Health 2008;85(5):677–92.

3. Sheehan TJ, Gershman ST, MacDougall LA, et al. Geographic assessment of breast cancer screening by towns, zip codes, and census tracts. J Public Health Manag Pract 2000;6(6):48–57.

4. King SJ, Livingston PM, Turner L, et al. Mapping the information resources available to patients with colorectal cancer. Eur J Cancer Care (Engl) 2010;19(4):492–500.

5. Mobley LR, Kuo T-M, Urato M, Subramanian S. Community contextual predictors of endoscopic colorectal cancer screening in the USA: spatial multilevel regression analysis. Int J Health Geogr 2010;9:44.

6. Rahman S, Price JH, Dignan M, Rahman S, Lindquist PS, Jordan TR. Access to Mammography Facilities and Detection of Breast Cancer by Screening Mammography: A GIS Approach. Int J Canc Prev 2009;2(6):403–13.

7. Maheswaran R, Pearson T, Jordan H, Black D. Socioeconomic deprivation, travel distance, location of service, and uptake of breast cancer screening in North Derbyshire, UK. J Epidemiol Community Health 2006;60(3):208–12.

8. Gwede CK, Ward BG, Luque JS, et al. Application of Geographic Information Systems and Asset Mapping to Facilitate Identification of Colorectal Cancer Screening Resources. Online J Public Health Inform 2010 [cited 2016 Jan 18];2(1). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3615755/

9. Maryland General Assembly. Senate Bill 234 . 2012;Available from: http://mlis.state.md.us/2012rs/billfile/sb0234.htm.

10. U.S. Food and Drug Administration,. Mammography Facilities . Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfMQSA/mqsa.cfm

11. Steffen, Ben. Maryland Ambulatory Surgery Provider Directory: 2014 Edition . Maryland Health Care Commission; 2015. Available from: http://mhcc.maryland.gov/mhcc/pages/hcfs/hcfs_amsurg/documents/con_ambsurg_directory_report_2014.pdf

12. Baltimore Neighborhood Indicators Alliance–Jacob France Institute. Vital Signs 10: Measuring Baltimore's Continued Progress Towards Strong Neighborhoods and a Thriving City. Baltimore, MD: University of Baltimore; 2012.

13. Ord JK, Getis A. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 1995;27(4):286–306.

14. Anselin L. Local Indicators of Spatial Association—LISA. Geographical Analysis 1995;27(2):93–115.

15. Moraga P, Montes F. Detection of spatial disease clusters with LISA functions. Stat Med 2011;30(10):1057–71.

16. ESRI. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute; 2011.

Tables and Figures:

Figure 4.1: Accredited cancer screening and treatment facilities in Baltimore City, Maryland

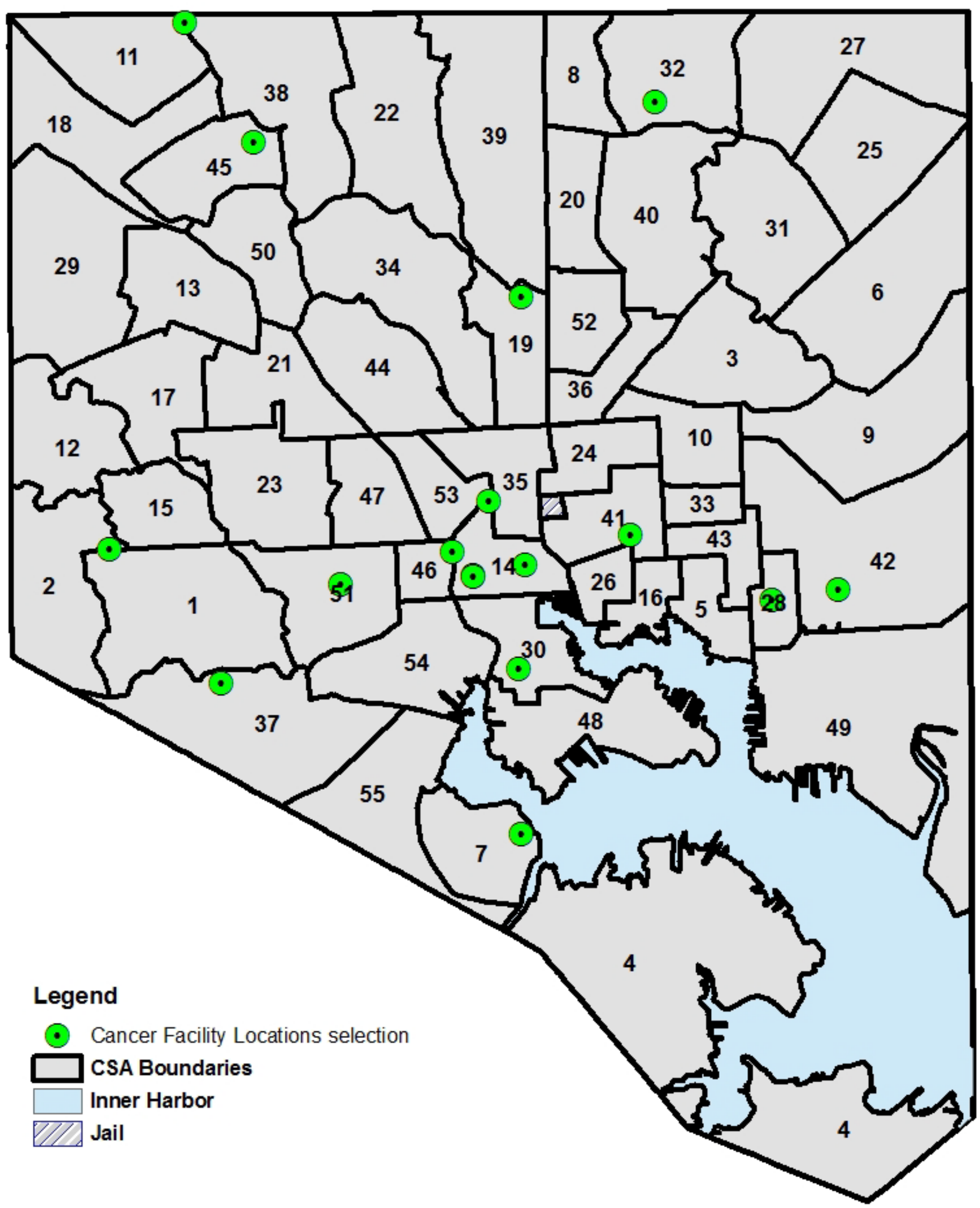


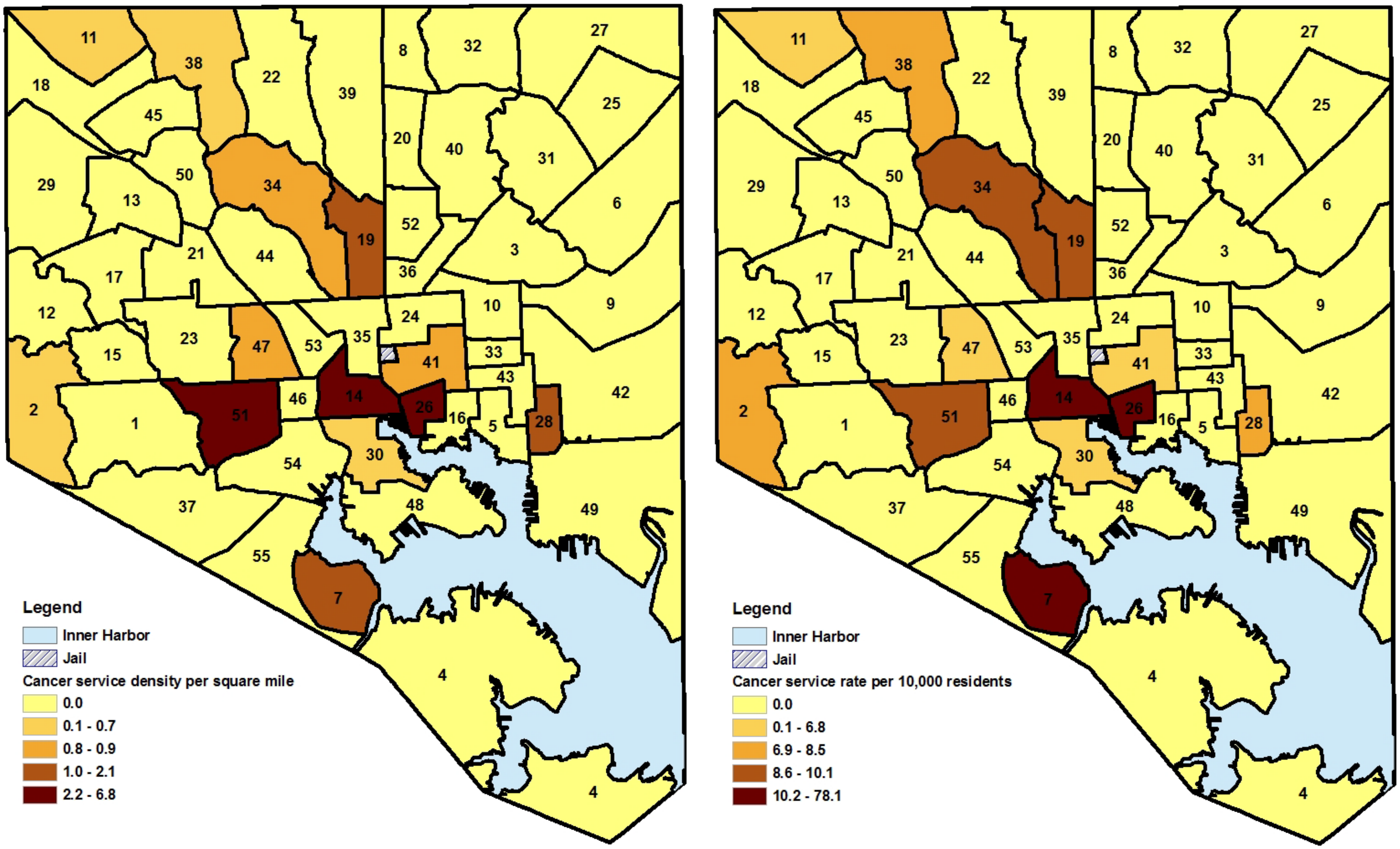
Table 4.1: Cancer facility addresses and services provided

Name	Address	Services provided			
		Breast cancer screening	Colorectal cancer screening	Breast cancer treatment	Colorectal cancer treatment
American Radiology Service	3700 Fleet St	Y	N	N	N
Bon Secours Hospital Outpatient	2000 W Baltimore St	Y	N	N	N
Downtown Baltimore Surgery Center, LLC	20 W West St	N	Y	N	N
Heritage Crossing	312 Martin Luther King Jr. Blvd	Y	N	N	N
Herman and Walter Samuels	2700 Quarry Lake Dr	Y	N	N	N
Johns Hopkins Bayview Medical Center	4940 Eastern Ave	Y	Y	Y	Y
Johns Hopkins Hospital, The Sidney Kimmel Comprehensive Cancer Center	401 N Broadway St	Y	Y	Y	Y
MedStar Good Samaritan Hospital	5601 Loch Raven Blvd	Y	Y	Y	Y
MedStar Harbor Hospital	3001 S Hanover St	Y	Y	Y	Y
MedStar Union Memorial Hospital	201 E University Pkwy	Y	Y	Y	Y
Mercy Medical Center - Baltimore, The Institute for Cancer Care	227 St Paul Pl	Y	Y	Y	Y
Saint Agnes Cancer Institute	900 S Caton Ave	Y	Y	Y	Y
Sinai Hospital of Baltimore, The Alvin & Lois Lapidus Cancer Institute (LifeBridge Health Network)	2401 W Belvedere Ave	Y	Y	Y	Y
University Care at Edmondson Village	4538 Edmondson Ave	Y	N	N	N
University of Maryland Greenebaum Cancer Center Fairway to Life Breast Care Center	22 S Greene St	Y	Y	Y	Y
University of Maryland Medical Center Midtown	827 Linden Ave	Y	N	N	N
VA Maryland Health Care System-Baltimore Division	10 North Greene St	Y	Y	Y	Y

Table 2.4: Key for Community Statistical Area (CSA) map

Allendale/Irvington/S. Hilton	1	Howard Park/West Arlington	29
Beechfield/Ten Hills/West	2	Inner Harbor/Federal Hill	30
Belair-Edison	3	Lauraville	31
Brooklyn/Curtis Bay	4	Loch Raven	32
Canton	5	Madison/East End	33
Cedonia/Frankford	6	Medfield/Hampden/Woodberry/Remington	34
Cherry Hill	7	Midtown	35
Chinquapin Park/Belvedere	8	Midway/Coldstream	36
Claremont/Armistead	9	Morrell Parkk/Violetville	37
Clifton-Berea	10	Mount Washington/Coldspring	38
Cross-Country/Cheswolde	11	North Baltimore/Guilford/Homeland	39
Dickeyville/Franklintown	12	Northwood	40
Dorchester/Ashburton	13	Oldtown/Middle East	41
Downtown/Seton Hill	14	Orangeville/E. Highlandtown	42
Edmonson Village	15	Patterson Park North & East	43
Fells Point	16	Penn North/Reservoir Hill	44
Forest Park/Walbrook	17	Pimlico/Arlington/Hilltop	45
Glen-Fallstaff	18	Poppleton/The Terraces/Hollins Market	46
Greater Charles Village/Barclay	19	Sandtown-Winchester/Harlem Park	47
Greater Govans	20	South Baltimore	48
Greater Mondawmin	21	Southeastern	49
Greater Roland Park/Poplar Hill	22	Southern Park Heights	50
Greater Rosemont	23	Southwest Baltimore	51
Greenmount East	24	The Waverlies	52
Hamilton	25	Upton/Druid Heights	53
Harbor East/Little Italy	26	Washington Village/Pigtown	54
Harford/Echodale	27	Westport/Mount Winans/Lakeland	55
Highlandtown	28		

Figure 4.2a-b: Cancer service density (facilities per square mile) and cancer service rate (facilities per 10,000 female residents aged 50 to 74 years) by quantile



Figures 4.3a-b: Hotspot analysis of cancer service density (facilities per square mile) and service rate (facilities per 10,000 female residents aged 50 to 74 years)

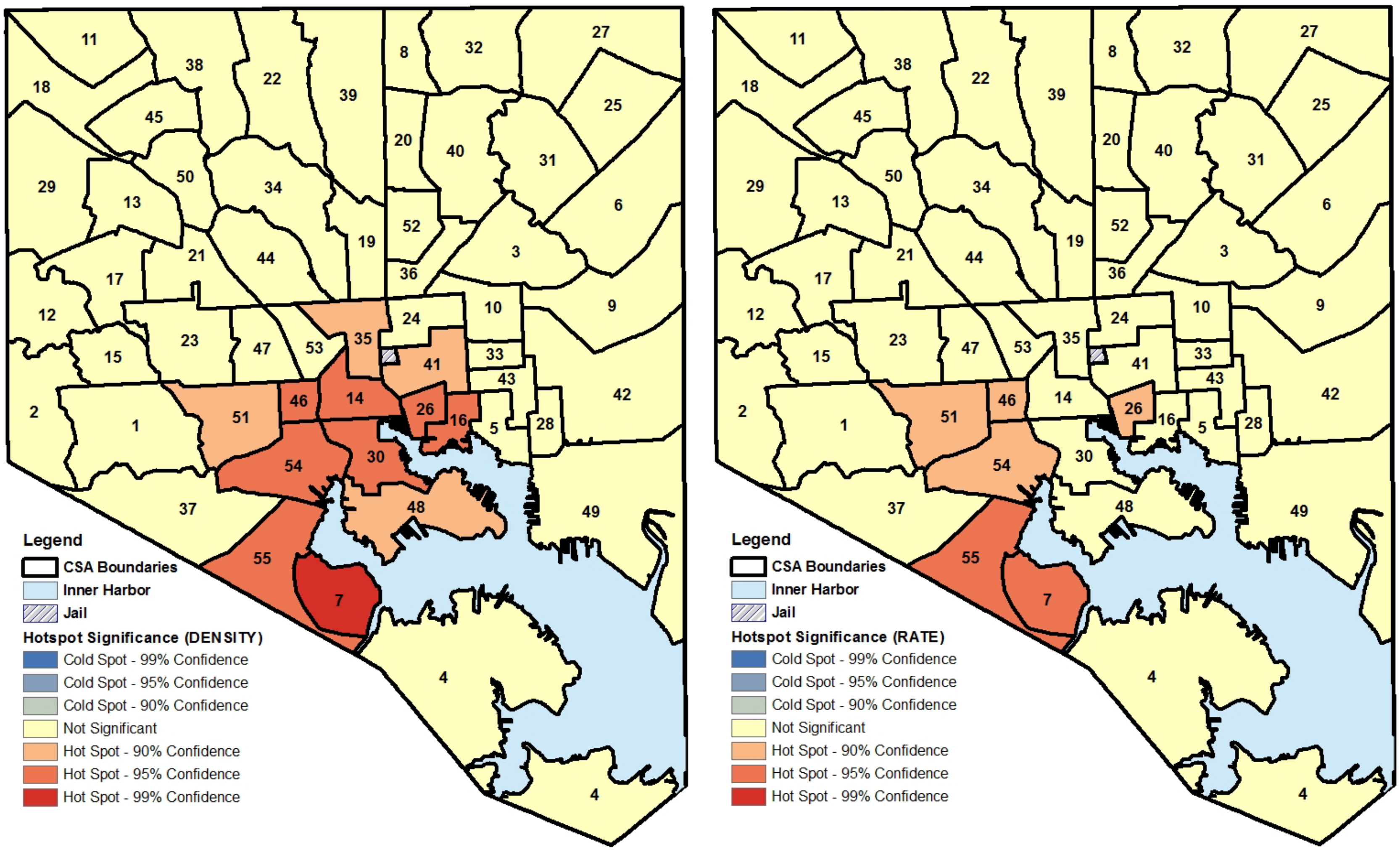
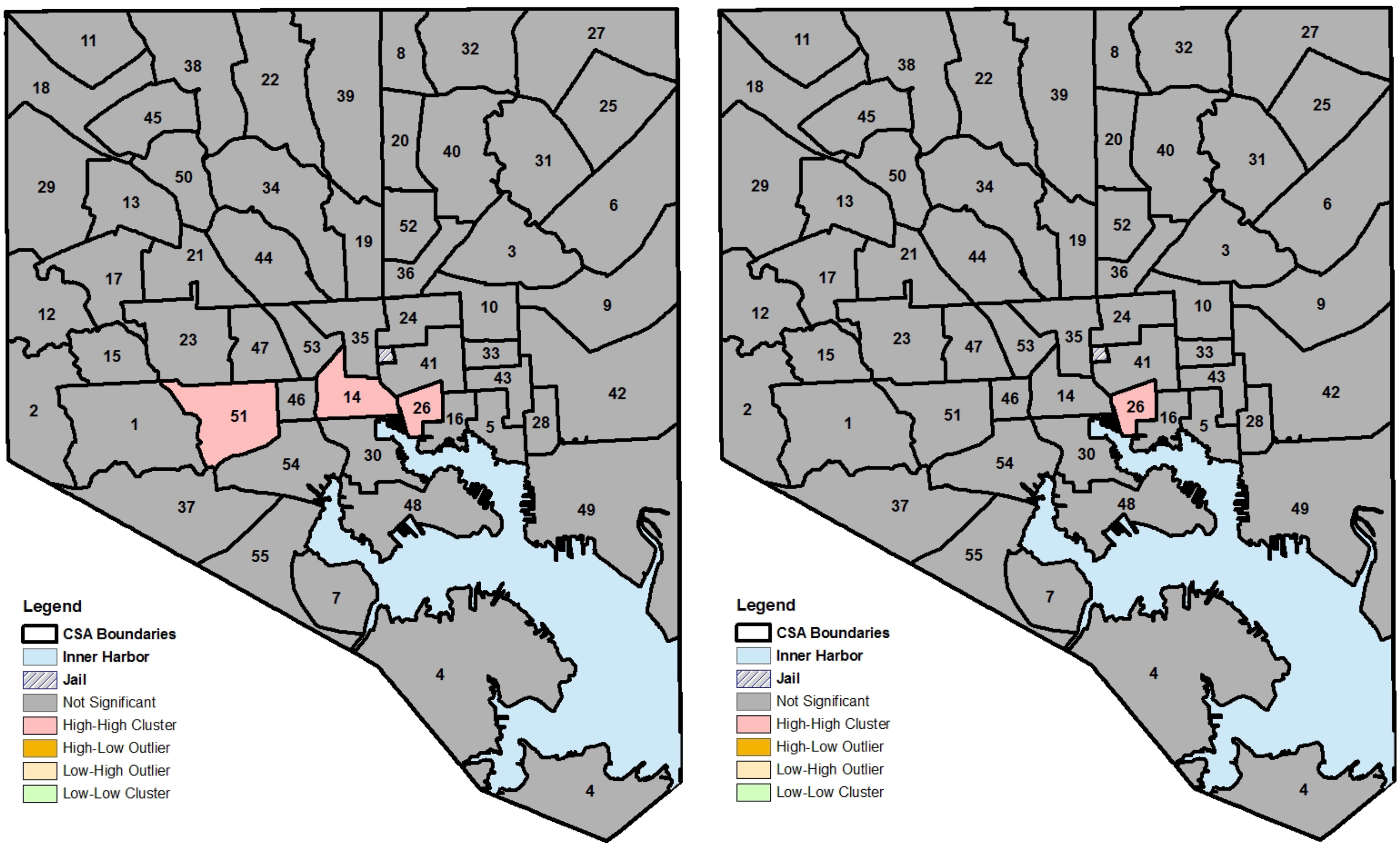


Table 4.3: Unadjusted and adjusted Ordinary Least Squares regression models for cancer service density and service rate with candidate neighborhood-level covariates

	Cancer service rate			Cancer service density	
	Coefficient	**p-value**		**Coefficient**	**p-value**
Simple models					
Females 50 to 74 years	0.001	0.489		-0.004	0.173
% African American	-0.004	0.413		-0.041	0.389
Racial Diversity Index	0.015	0.038*		0.143	0.059
Household income <25K	0.013	0.294		0.079	0.530
Female headed	0.009	0.249		0.062	0.460
Vacants	0.011	0.467		0.013	0.936
Housing violations	0.059	0.448		0.313	0.702
Crime	0.013	<0.001*		0.151	<0.001*
Domestic violence	0.016	0.104		0.142	0.174
Teen births	0.000	0.981		-0.013	0.782
Employed	-0.015	0.277		-0.070	0.622
Businesses	0.002	<0.001*		0.019	<0.001*
Voted	-0.030	0.075		-0.221	0.211
Dirty streets	0.000	0.908		-0.017	0.535
Tree coverage	-0.015	0.092		-0.108	0.254
Neighborhood associations	-0.006	0.835		-0.165	0.577
Adjusted					
Racial Diversity Index	-0.001	0.742	Crime	0.084	<0.001*
Crime	0.006	0.032*	Businesses	0.010	0.003*
Businesses	0.001	0.010*			
R-squared	0.716			0.817	

* Statistically significant

Figures 4.4a-b: Local Moran's I analysis for cancer service density (facilities per square mile) and service rate (facilities per 10,000 female residents aged 50 to 74 years)



CHAPTER 5: SUMMARY AND CONCLUSIONS

Key findings

The physical and social conditions in which people reside have become topics of discussion when considering the likely drivers of cancer disparities, especially those observed when outcomes vary geographically. While there has been a growing movement towards accounting for hierarchical effects through multi-level models, there has still been a disconnect in identifying clusters with a high public health burden and leveraging those findings into practical strategies in health policy. This dissertation has evaluated the impact of unique neighborhood-level characteristics on several cancer-related outcomes, which include incidence, mortality, and the locations of cancer facilities.

Existence of geographic variation across cancer-related outcomes

For all three aims, the primary objective was to determine whether the cancer-related outcomes of interest exhibited geographic variation within Baltimore City. Upon completion of each analysis, it became readily apparently that cancer incidence (Aim 1), cancer mortality (Aim 2), and physical access (Aim 3) to facilities were not randomly distributed and displayed a geographic distribution where the outcomes were concentrated in specific sections of the city. These findings confirmed the dissertation's overarching hypothesis that the spatial patterns displayed by health-related outcomes are not random. The mechanisms that drive this aggregation go beyond the traditional etiological pathways. The social infrastructure of these neighborhoods plays a role in facilitating the differential clustering of various outcomes across the disease spectrum.

Heterogeneity in geographic variability across cancer types

While it was hypothesized that the primary endpoints of interest would exhibit clustering at a scale more granular than county-level (i.e., Community Statistical Area), the degree to which the clustering varied by cancer site for the same outcome was unexpected. While there were instances where the concentration of events for one cancer overlapped with those of another, the more common scenario involved cancer type playing the role of effect modifier in the observed geographic distribution. For example, breast cancer incidence had an aggregation of significantly high incidence in the northern area of Baltimore City while displaying low incidence in the downtown area around Inner Harbor. However, cervical cancer had a CSA with significantly high incidence but had also yielded low rates for breast. This scenario occurred in several instances across each of the analyses. The data provides evidence that caution should be heeded when allocating programs to specific neighborhoods that address different disease types. It could very well be the case that a community's risk profile for a specific cancer is drastically different than its risk profile for another tumor type.

Heterogeneity in associations with neighborhood-level characteristics

In each of the analyses, the association between neighborhood characteristics and the geographic distribution for each of the cancer outcomes was evaluated through site-specific ordinary least squares (OLS) regression models. These methods yielded the unexpected finding that the community-level characteristics significantly associated with those distributions varied by cancer site. This would indicate that the neighborhood mechanisms likely driving the observed geographic variations are cancer-site specific. It should be noted that any significant associations observed between community-level data and outcome data, which was collected at the individual-level and aggregated to the area-level, should be interpreted while keeping the ecological model in mind. In other words, aggregate data should not be used to make inferences at the individual-level, which would result ecologic fallacy.

Implications and future directions

While innovations have assisted in significantly improving cancer outcomes over the last decade, this progress has not uniformly benefited subpopulations, particularly those that have traditionally been underserved. The integration of spatial statistics into cancer disparities research has been underutilized in addressing this gap. Previously, there has been the common approach of focusing on only etiological factors and their unequal distribution in specific subgroups. However, this framework does not account for the social conditions in which these etiological associations are occurring. It has been demonstrated that neighborhood-level factors have an effect on health that is independent of individual-level covariates.[1,2] The context in which disease occurs could very well modify many of the well-known exposure-outcome relationships.

Research implications

The results yielded by the analyses demonstrated a noticeable variation in how cancer outcomes were geographically distributed across Baltimore City. This finding aligned with the literature that has recently delved into the geographic variation of cancer outcomes. Studies have demonstrated that cancer-related outcomes do display spatial patterns that correlate with aggregate-level characteristics.[3-7] However, the more significant findings that add new information to the current knowledge consisted of: 1) variation existing within cancer outcomes across cancer sites; and 2) the neighborhood covariates associated with the distribution varied by cancer site as well. This information is relevant from a research perspective because it highlights the need to conduct cancer site-specific analyses. By treating all primary sites the same within a regression model or spatial statistics, initial findings might yield null results and lead to a generalization that there is no significant association of interest present.

An additional implication is the importance of the unit of analysis utilized in geospatial assessments of exposures and outcomes. Currently, many updates on cancer control in Maryland are provided at the county-level. Each of the chapters demonstrated that a single summary measure describing Baltimore City's incidence, mortality, and facility concentration would have masked the pockets of burden occurring at the level of the CSA. There is an over-reliance on the global magnitude of disease, which takes away from the importance of understanding small area estimates in addressing geographic cancer disparities.

Also, it should be noted that while the methods as they are currently described could provide considerable leverage in furthering cancer disparities research, this could have been further maximized had it not been missing data. The treatment information provided by the Maryland Cancer Registry (MCR) was limited in its breadth and depth. The organization's scarce resources make it difficult to follow-up with each of the cancer cases and update the course of treatment initiated post-diagnosis. The degree of missingness for this particular covariate prevented the research opportunity of understanding whether treatment utilization and timeliness varied geographically. This area, in particular, would have provided a potential avenue for intervening on a factor that is significantly associated with prognosis and survival.

Implications for policy

Beyond the research implications discussed, there are several actionable policy items that could make use of the results in very practical ways. As alluded to from a research perspective, resources need to be prioritized and dedicated to capturing the appropriate data to help inform decision-making. From a policy stance, there is a need to acknowledge that a group such as the MCR needs added capacity to appropriately collect data that will help more efficiently meet the cancer control goals set out to improve rates and outcomes. In a similar manner, the geospatial information generated should be utilized in a meaningful way. There has been a slight shift that has begun to take this approach. This has been illustrated through the passing of the Maryland Health Improvement and Disparities Reduction Act of 2012.[8] This particular bill aims to establish Health Enterprise Zones in areas that are under-resourced. It seems that mapping disease burden and the current allocation resources to address it is a practical first step when making decisions.

It is also important to recognize how this dissertation fits within the scope of the Affordable Care Act, specifically regarding the new requirements that were put in place for organizations to retain their tax-exempt status.[9] As part of the regulations, 501(c)(3) organizations that oversee at least one hospital facility must have each of its facilities carry out a community health needs assessment (CHNA) and subsequently implement a strategy at least every three years. A failure to follow through on these requirements would result in an excise tax. In line with the previous discussion of mapping disease burden, it is critical that hospitals and clinics evaluate the needs of their patient population. Oftentimes, change takes effect more quickly when there is a monetary penalty or cost incentive tied to it. This dissertation provides another method to assess community needs and leveraging that information in a way that can direct the efficient allocation of resources.

Additionally, the comprehensiveness of the neighborhood-level data collected in Baltimore City is what helped to drive the uniqueness of this dissertation. There was a local commitment to monitor the quality of life for each of the city's neighborhoods. The effort to collect these measures provided a new evaluation of the relationship between the social conditions and cancer outcomes. The previous literature primarily utilized the information captured by the U.S. Census, which is limited in its scope and is collected less frequently. Policymakers outside of Baltimore City need to understand that more informative and evidence-based decisions can be made if there is a deeper understanding of the conditions of their communities through rigorously collected data.

Finally, the methodology presented took a transdisciplinary approach, which is often underutilized in studies that focus mostly on clinical associations. It is necessary to expand findings beyond just the academic setting. The motivations for research should be considered through the lens of impactful policy implications. The estimates yielded by this dissertation's results improve statistical precision while also maintaining geographic resolution to demonstrate small-scale patterns within neighborhoods but also large-scale trends across the whole county. The overarching implication is that it can assist in laying the groundwork for future geostatistical analyses that direct allocation of cancer prevention/treatment services in a more effective way and to develop geographically tailored interventions, which has been accomplished in other communities.[10-12] "To win the war against cancer, we must apply what we know *at any given time* to all people."[13]

References:

1. Warnecke RB, Oh A, Breen N, et al. Approaching health disparities from a population perspective: the National Institutes of Health Centers for Population Health and Health Disparities. Am J Public Health 2008;98(9):1608–15. 12.

2. Gehlert S, Colditz GA. Cancer disparities: unmet challenges in the elimination of disparities. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol 2011;20(9):1809–14.

3. Rushton G, Peleg I, Banerjee A, Smith G, West M. Analyzing geographic patterns of disease incidence: rates of late-stage colorectal cancer in Iowa. J Med Syst 2004;28(3):223–36.

4. McElroy JA, Remington PL, Gangnon RE, Hariharan L, Andersen LD. Identifying geographic disparities in the early detection of breast cancer using a geographic information system. Prev Chronic Dis 2006;3(1):A10.

5. Singh GK, Miller BA, Hankey BF, Edwards BK. Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975–1999. Bethesda, MD: National Cancer Institute; 2003.

6. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. Br J Cancer 2002;87(6):585–90.

7. Klassen AC, Pankiewicz A, Hsieh S, Ward A, Curriero FC. The association of area-level social class and tobacco use with adverse breast cancer characteristics among white and black women: evidence from Maryland, 1992-2003. Int J Health Geogr 2015;14:13.

8. Maryland General Assembly. Senate Bill 234. 2012;Available from: http://mlis.state.md.us/2012rs/billfile/sb0234.htm.

9. Internal Revenue Service. "New Requirements for 501(c)(3) Hospitals Under Affordable Care Act." IRS. Updated 9 June 2015. Available from: https://www.irs.gov/Charities-&-Non-Profits/Charitable-Organizations/New-Requirements-for-501(c)(3)-Hospitals-Under-the-Affordable-Care-Act

10. Alcaraz KI, Kreuter MW, Bryan RP. Use of GIS to identify optimal settings for cancer prevention and control in African American communities. Prev Med 2009;49(1):54–7.

11. Brantley-Sieders DM, Fan K-H, Deming-Halverson SL, Shyr Y, Cook RS. Local breast cancer spatial patterning: a tool for community health resource allocation to address local disparities in breast cancer mortality. PloS One 2012;7(9):e45238.

12. King SJ, Livingston PM, Turner L, et al. Mapping the information resources available to patients with colorectal cancer. Eur J Cancer Care (Engl) 2010;19(4):492–500.

13. Freeman, HP. Poverty, Culture, and Social Injustice: Determinants of Cancer Disparities. CA: A Cancer Journal for Clinicians 2004; 54(2):72-77.

Curriculum Vitae
Aracelis Torres

Personal Data

Phone: 917-941-3726
Email: aracelis.z.torres@gmail.com

Education and Training

2011 – present	**Johns Hopkins School of Public Health, Baltimore, MD** Doctor of Philosophy (PhD) candidate Department of Epidemiology Dissertation topic: *Evaluating geographic cancer disparities in Baltimore City, Maryland*
2009 – 2011	**Yale University School of Public Health, New Haven, CT** Master of Public Health (MPH) Department of Chronic Disease Epidemiology Thesis topic: *Mammography Screening in Latino Women Living in the Northeast US: Qualitative Research Findings*
2004 – 2008	**Yale University, New Haven, CT** Bachelor of Science Molecular, Cellular, and Developmental Biology

Professional Experience

2015 – present	**Quantitative Scientist**, Science Flatiron Health, New York, NY
2015	**Research Assistant**, Health Situation Room Brooklyn Hospital Center, Brooklyn, NY
2013	**Professional Epidemiology Research Assistant**, Office of Epidemiologic Services Baltimore City Health Department, Baltimore, MD
2012	**Graduate Student Investigator**, Center to Reduce Cancer Disparities Johns Hopkins School of Public Health, Baltimore, MD
2010 – 2011	**Research Associate**, Center for Cancer Epidemiology and Prevention Yale University School of Public Health, New Haven, CT
2008 – 2009	**Special Medical Systems Analyst**, Chronic Disease Management William F. Ryan Community Health Center, New York, NY

Honors and Awards

2011 – present	The C. Sylvia and Eddie C. Brown Community Health Scholarship
2014	Harvey M. Meyerhoff Fellowship in Cancer Prevention
2014	Jean Coombs Fund
2013	Miriam E. Brailey Fund
2012	AACR Scholar-in-Training Award

Presentations and Posters

2012	**Distance to Mammography Facilities Modifies the Effect of Patient Navigation on Breast Cancer Screening Adherence among Female Black Medicare Beneficiaries in Baltimore City** Poster, 5th American Association for Cancer Research (AACR) Conference on the Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved
2010	**Mammography Screening in Latino Women Living in the Northeast US: Qualitative Research Findings** Poster, 138th American Public Health Association (APHA) Annual Meeting Prevention of Disease in Latino Communities Session

Teaching

2014	**Co-instructor**, Epidemiology Workshop: Interpreting and Using Epidemiological Evidence Secretariat of Health Surveillance, Ministry of Health Brasília- DF, Brazil
2012 – 2013	**Teaching Assistant**, Epidemiologic Methods I & II Department of Epidemiology Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
2010	**Teaching Fellow**, Principles of Epidemiology I Chronic Disease Epidemiology Department Yale University School of Public Health, New Haven, CT